August 16, 2018

Comments from Academics, Scientists and Clinicians on Problem Formulations for the Risk Evaluations to Be Conducted Under the Toxic Substances Control Act (TSCA)

Submitted online via *Regulations.gov* to dockets EPA–HQ–OPPT–2016–0736, EPA–HQ–OPPT–2016–0741, EPA–HQ–OPPT–2016–0723, EPA–HQ–OPPT–2016–0733, EPA–HQ–OPPT–2016–0735, EPA–HQ–OPPT–2016–0742, EPA–HQ–OPPT–2016–0743, EPA–HQ–OPPT–2016–0725, EPA–HQ–OPPT–2016–0732 and EPA–HQ–OPPT–2016–0737

These comments are submitted on behalf of the undersigned academics, scientists, and clinicians. We declare collectively that we have no direct or indirect financial or fiduciary interest in any chemical or product that is the subject of these comments. The co-signers' institutional affiliations are included for identification purposes only and do not imply institutional endorsement or support unless indicated otherwise.

We appreciate the opportunity to provide written comments on the Environmental Protection Agency's (EPA) problem formulations for the risk evaluations of the first 10 TSCA chemicals. These chemicals have a collective aggregate production volume of more than 1 billion pounds a year,[1] and people encounter them routinely in their daily lives—such as in dry cleaning, personal care products, electronics and building materials.

With these evaluations and their subsequent implications for risk management decisions, EPA should be protecting the public, including sensitive groups, from chemicals that pose an unreasonable risk as required by law. However, we have serious concerns about the approach EPA outlines in its problem formulations, which is inconsistent both with TSCA mandates and current scientific principles and methods to assess chemical risks. One fundamental flaw is that EPA proposes excluding populations with known exposures to these chemicals from most of its risk evaluations. Potentially exposed and susceptible sub-populations such as pregnant women, infants and children are being excluded from the risk evaluations, in conflict with EPA's statutory obligations and best scientific practice. Further, even for the populations EPA has chosen to assess, it will exclude significant known sources of exposures. EPA is thus assessing only a minor proportion of likely exposures.

TSCA statute[2] and regulation[3] require that EPA has adequate data on chemicals to inform its risk evaluations. EPA has the authority to request needed toxicity and/or exposure testing,[4] yet is choosing not to request additional data for any of the 10 chemicals.[5] In particular, EPA does not have adequate data on pigment violet 29 to make a risk determination.

[1] This is the aggregate production volume estimate for the 9 chemicals with production volume information available. Asbestos importers claimed production volumes as confidential business information (CBI).
[2] 15 USC §2601 (b)(1)
[3] 40 CFR § 702.41 (b)
[4] 15 USC §2603
[5] US EPA (May 2018) EPA's Responses to Public Comments Received on the Scope Documents for the First Ten

Overall, EPA's approach to the problem formulations is not consistent with TSCA requirements, is not scientifically supported, and will lead to inaccurate evaluations that substantially underestimate risks—with the ultimate result of missing or insufficient limits on chemicals that could pose unreasonable risks to the public's health. EPA must request testing so it has adequate data and evaluate a full set of exposures from all conditions of use as detailed below.

Our comments address the following main points:

1. **The problem formulations do not consider the following exposures and will result in significant underestimates of actual risk, including for potentially exposed or susceptible sub-populations.**
    a. **EPA must consider aggregate exposure within and across all exposed populations.**
    b. **EPA must consider exposures from legacy uses and uses where a chemical is present as a contaminant or by-product in the exposure assessments for all populations.**
    c. **EPA must include chemical exposures from air, water, land and all other pathways in the exposure assessment for all populations, regardless of claims of coverage under other environmental statutes.**
    d. **EPA must assess risks to the general population for all 10 chemicals, and specifically to groups in proximity to conditions of use as a potentially exposed sub-population. Existing risks to the general population will be ignored for the following chemicals, as the problem formulations propose to exclude the general population entirely: PERC; asbestos; TCE; NMP; methylene chloride; carbon tetrachloride; 1, 4 dioxane; and pigment violet 29.**
    e. **EPA must use realistic occupational exposure scenarios and not assume the use of exposure controls or compliance with exposure standards.**
2. **EPA should identify potential susceptible sub-populations based on established, scientifically supported extrinsic and intrinsic factors that increase vulnerability.**
3. **EPA should request additional test data for pigment violet 29 as the available data does not meet the TSCA requirement of adequate information to make a risk determination.**
4. **EPA should move forward with finalizing rules to limit uses of TCE and NMP immediately. EPA should not re-evaluate conditions of use of TCE and NMP already found to pose unreasonable risks.**

We are appreciative of the opportunity to provide public input. Please do not hesitate to contact us with any questions regarding these comments.

Sincerely,

Veena Singla, PhD
Associate Director, Science and Policy, Program on Reproductive Health and the Environment
University of California, San Francisco

Tracey Woodruff, PhD, MPH
Professor, Department of Obstetrics, Gynecology and Reproductive Sciences
Director, Program on Reproductive Health and the Environment
University of California, San Francisco

Chemicals for Risk Evaluation under TSCA. Pg. 10-11 "As of now, EPA has not identified the need for any such testing for the first 10 chemicals." Available: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/response-comments-issues-impacting-first-10-chemicals

Juleen Lam, PhD, MHS, MS
Assistant Professor, Department of Health Sciences
California State University, East Bay

Patricia D. Koman, PhD, MPP
President and Senior Health Scientist
Green Barn Research Associates*

Liz Borkowski, MPH
Senior Research Scientist, Milken Institute School of Public Health
George Washington University

Sheila Brear, BDS
Associate Dean, Academic Affairs, School of Dentistry
University of California, San Francisco

Phil Brown, PhD
Adjunct Professor of Pathology and Laboratory Medicine
Brown University

Adelita G. Cantu, PhD, RN
Associate Professor
Alliance of Nurses for Healthy Environments

Danielle Fries, MPH
Science Associate, Program on Reproductive Health and the Environment
University of California, San Francisco

Robert M. Gould, MD
Associate Adjunct Professor, Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco
Past-President, Physicians for Social Responsibility

Jyotsna Jagai, MS, MPH, PhD
Research Assistant Professor
University of Illinois at Chicago

Jean-Marie Kauth, PhD, MPH
Professor
Benedictine University

Janet Pelrman, MD, MPH
Physician
Stanford Childrens Health

Ted Schettler MD, MPH
Science Director
Science and Environmental Health Network

Rachel M. Shaffer, MPH
PhD Candidate
UW Seattle School of Public Health

Patrice Sutton, MPH
Research Scientist, Program on Reproductive Health and the Environment
University of California, San Francisco

Marya Zlatnik, MD
Clinical Professor, Department of Obstetrics, Gynecology & Reproductive Sciences
University of California, San Francisco

*indicates organizational support

DETAILED COMMENTS

1. **The problem formulations do not consider the following exposures and will result in significant underestimates of actual risk, including for potentially exposed or susceptible sub-populations.**

1. **(a) EPA must consider aggregate exposure within and across all exposed populations.**

EPA is proposing to exclude known sources of chemical exposure including air, water, and soil for all 10 chemical evaluations.[6] This plan does not meet the TSCA requirement for EPA to determine whether "the manufacture, processing, distribution in commerce, use, or disposal of a chemical substance or mixture, or that any combination of such activities, presents an unreasonable risk of injury to health or the environment," including to potentially exposed or susceptible sub-populations.[7] The plan is also not consistent with TSCA's mandate to use the "best available science"[8] because it ignores established scientific principles for exposure assessment.[9,10] To meet TSCA mandates, EPA must consider total chemical exposure-- the aggregate exposure--for all exposed populations.

Risk evaluations based on these proposals would consistently underestimate the true risk to the public. EPA has described the concept of assessing aggregate exposures as "the risk cup," where every use of a chemical contributes to filling the cup.[11] The Agency can only determine if risks exceed levels of concern, that is whether the risk cup is full or overflowing, by adding together all contributing exposures. However, if known chemical uses and exposures are ignored, the cup levels will be an underestimate of the true risk posed, suggesting that risks are below levels of concern when in reality the cup might be full or overflowing, indicating an unreasonable risk that warrants action.

Accurate assessment of aggregate exposure may also reveal risks to susceptible sub-populations that would have been missed if only limited exposure sources were considered. For example, EPA's 2005 risk assessment of the pesticide sulfuryl fluoride found that "Although sulfuryl fluoride residues in food contribute only a very small portion of total exposure to fluoride, when combined with other fluoride exposure pathways, including drinking water and toothpaste, EPA has concluded that the tolerance (legal residue limits on food) no longer meets the safety standard…aggregate fluoride exposure is too

---

[6] The following paragraph is included in all 10 problem formulations: "As part of this problem formulation, EPA also identified exposure pathways under other environmental statutes, administered by EPA, which adequately assess and effectively manage exposures and for which long-standing regulatory and analytical processes already exist, i.e., the Clean Air Act (CAA), the Safe Drinking Water Act (SDWA), the Clean Water Act (CWA) and the Resource Conservation and Recovery Act (RCRA). OPPT worked closely with the offices within EPA that administer and implement the regulatory programs under these statutes. In some cases, EPA has determined that chemicals present in various media pathways (i.e., air, water, land) fall under the jurisdiction of existing regulatory programs and associated analytical processes carried out under other EPA-administered statutes and have been assessed and effectively managed under those programs. EPA believes that the TSCA risk evaluation should focus on those exposure pathways associated with TSCA uses that are not subject to the regulatory regimes discussed above because these pathways are likely to represent the greatest areas of concern to EPA."

[7] 15 USC §2605(b)

[8] 15 USC §2625(h)

[9] US EPA (1992) Guidelines for Exposure Assessment. Risk Assessment Forum, Washington DC. Available: https://www.epa.gov/sites/production/files/2014-11/documents/guidelines_exp_assessment.pdf

[10] National Academies of Sciences, Engineering, and Medicine (2017) *Using 21st Century Science to Improve Risk-Related Evaluations*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24635.

[11] US EPA (January 31, 1997) PRN 97-1: Agency Actions under the Requirements of the Food Quality Protection Act. Available: https://www.epa.gov/pesticide-registration/prn-97-1-agency-actions-under-requirements-food-quality-protection-act#risk

high for certain identifiable subpopulations in the United States, in particular children under the age of 7…"[12] Based on this finding, EPA proposed a rule to restrict sulfuryl fluoride uses on food.[13]

The risk assessment for sulfuryl fluoride considered the parent chemical and its breakdown product, fluoride. EPA assessed aggregate exposure to fluoride by adding together exposures from all sources, including food, beverages, drinking water and consumer products (toothpaste).[14] Even though fluoride had existing assessments, regulations and standards under other statutes administered by EPA (the Safe Drinking Water Act[15]) and other agencies (the Food and Drug Administration for fluoride in toothpaste), it was critical that EPA included those sources in its risk assessment to accurately assess total exposure, and thus real-world risk, from fluoride. With the aggregate exposure assessment, EPA found that "Most people in the United States are not exposed to unsafe levels of fluoride. However, aggregate fluoride exposure for infants and children under the age of 7 years old, where drinking water contains high levels of natural fluoride, exceeds the level that can cause severe dental fluorosis."[16] If EPA had only considered the risk from fluoride residues contributed by sulfuryl fluoride in isolation, its assessment would not have identified the existing risks to infants and children. This demonstrates the importance of considering all potential exposures in an aggregate exposure assessment as well as the Agency's ability to conduct such aggregate evaluations.

Yet, in the problem formulations for all 10 chemicals, looking at single chemical exposures in isolation is exactly what EPA is proposing to do, resulting in EPA systematically underestimating exposures, leaving the public and susceptible sub-populations at risk. TSCA mandates that EPA conduct risk evaluations to determine whether a chemical substance presents an unreasonable risk to a potentially exposed or susceptible sub-population. As the sulfuryl fluoride example shows, EPA cannot meet this mandate without considering aggregate exposures. The Agency has decades of experience conducting risk assessments that assess aggregate exposures for pesticides, air contaminants and other chemicals and has developed methods to estimate exposures for routes or pathways with limited data. This includes using science-based defaults for areas where there is missing data, an approach that has been recommended by the National Academy of Sciences.[17] There are no scientific or technical barriers for EPA to conduct an aggregate exposure assessment and the methods are established in Agency practice.

In summary, EPA needs to account for all the sources of exposure or it will underestimate risk for all 10 chemicals. Because "sentinel exposure" does not consider all sources of exposure, it should only be included **as part of** assessing aggregate exposure-- these should not be considered mutually exclusive and EPA should not choose to incorporate sentinel over aggregate. Instead, we recommend that sentinel exposure be considered within the context of aggregate exposure when appropriate.

---

[12] US EPA. EPA Proposes to Withdraw Sulfuryl Fluoride Tolerances. Available: https://archive.epa.gov/oppsrrd1/registration_review/web/html/evaluations.html

[13] 76 FR 3421 (January 19, 2011) Available: https://www.federalregister.gov/documents/2011/01/19/2011-917/sulfuryl-fluoride-proposed-order-granting-objections-to-tolerances-and-denying-request-for-a-stay

[14] US EPA (2005) Human Health Risk Assessment for Sulfuryl Fluoride and Fluoride Anion Addressing the Section 3 Registration of Sulfuryl Fluoride Fumigation of Food Processing Facilities. Available: https://www.regulations.gov/document?D=EPA-HQ-OPP-2005-0174-0009

[15] US EPA (2011) Questions and Answers on Fluoride. Available: https://www.epa.gov/sites/production/files/2015-10/documents/2011_fluoride_questionsanswers.pdf

[16] US EPA. EPA Proposes to Withdraw Sulfuryl Fluoride Tolerances. Available: https://archive.epa.gov/oppsrrd1/registration_review/web/html/evaluations.html

[17] National Research Council (2009) *Science and Decisions: Advancing Risk Assessment*. Washington, DC: The National Academies Press. https://doi.org/10.17226/12209.

1. (b) EPA must consider exposures from "legacy uses" and uses where a chemical is present as a contaminant or by-product in the exposure assessments for all populations.

EPA proposes to exclude exposures related to "legacy uses" and uses where a chemical is present as a contaminant or by-product for all 10 problem formulations. For example with asbestos, EPA states, "In the case of asbestos, legacy uses, associated disposals, and legacy disposals will be excluded from the problem formulation and risk evaluation… (t)hese include asbestos-containing materials that remain in older buildings or are part of older products but for which manufacture, processing and distribution in commerce are not currently intended, known or reasonably foreseen."[18]

As discussed above, this is not consistent with the best available science and will result in an inaccurate exposure assessment because though the use may not be ongoing, the exposures *are* ongoing—that is, though asbestos is no longer being used in some building materials, the existing asbestos in buildings still results in exposures. Thus "legacy uses" contribute to aggregate exposures and risks and must be considered in the risk evaluation.

As another example, EPA proposes to exclude from consideration conditions of use of 1,4-dioxane when it is present as contaminant in a wide variety of items, including household detergents, cosmetics/ toiletries, and foods.[19] This exclusion is not scientifically justified. Cosmetics and personal care products have the potential to contribute significantly to exposures, since people are applying them directly to their bodies, often multiple times per day, every day. Although 1,4-dioxane may not have been intentionally added to these products, its presence as a contaminant contributes to exposures to consumers using these products. Therefore, as discussed earlier, ignoring these exposures will result in underestimating existing risks to the population.

Lastly, for HBCD flame retardants, EPA states, "There is no longer manufacture, processing or distribution of HBCD for HIPS or textiles; and therefore, those uses are not included in the scope of the risk evaluation of HBCD."[20] HBCD was used as an additive flame retardant in high impact polystyrene (HIPS) casing for electronics such as TVs, DVD players, computers, etc. A recent study found a significant correlation between the number of electronics in the home and the amount of HBCD on people's hands (an exposure metric used to estimate dermal absorption and hand-to-mouth ingestion), indicating that electronics are a significant source of exposure for people.[21] Toddlers and young children, a potential susceptible sub-population, can have greater exposures to environmental chemicals compared to adults due to their behaviors (i.e., increased hand-to-mouth ingestion rates) and physiological differences.[22]

---

[18] US EPA (May 2018) Problem Formulation of the Risk Evaluation for Asbestos. Pg. 8

[19] US EPA (May 2018). Problem Formulation of the Risk Evaluation for 1,4-Dioxane. Pg. 18

[20] US EPA (May 2018) Problem Formulation for Cyclic Aliphatic Bromides Cluster (HBCD). Pg. 9

[21] Tay JH, Sellström U, Papadopoulou E, Padilla-Sánchez JA, Haug LS, de Wit CA. Assessment of dermal exposure to halogenated flame retardants: Comparison using direct measurements from hand wipes with an indirect estimation from settled dust concentrations. Environ Int. 2018 Jun 1;115:285–94.

[22] US EPA. Children Are Not Little Adults! Available: https://www.epa.gov/children/children-are-not-little-adults

Moya J, Bearer CF, Etzel RA. Children's behavior and physiology and how it affects exposure to environmental contaminants. Pediatrics. 2004 Apr 1;113(Supplement 3):996-1006.

Bearer, C.F., 1995. How are children different from adults? Environmental health perspectives, 103(Suppl 6), p.7.

Goldman, L.R., 1995. Children--unique and vulnerable. Environmental risks facing children and recommendations for response. Environmental Health Perspectives, 103(Suppl 6), p.13.

Further, studies find that dermal exposure from contact with HBCD-treated furniture and textiles contributes substantially to the human body burden of HBCD; a specific recommendation is that "Future risk assessments for these contaminants…should consider dermal contact with treated products as a potential significant human exposure pathway to these hazardous chemicals,"[23] yet EPA has made the decision to not include this. Excluding HIPS and textile uses of HBCD will result in a significant exposure underestimates, particularly for children.

EPA must include all uses that contribute or are reasonably foreseen contributions to exposures for all populations. Failure to do so will underestimate risk for all 10 chemicals.

1. **(c) EPA must include chemical exposures from air, water, land and all other pathways in the exposure assessment for all populations, regardless of claims of coverage under other environmental statutes.**

For all 10 problem formulations, EPA is proposing to exclude "exposure pathways under other environmental statutes, administered by EPA, which adequately assess and effectively manage exposures and for which long-standing regulatory and analytical processes already exist, i.e., the Clean Air Act (CAA), the Safe Drinking Water Act (SDWA), the Clean Water Act (CWA) and the Resource Conservation and Recovery Act (RCRA)."[24]

There are several scientific problems with this statement and approach:

- As discussed in 1(a) above, these pathways contribute to aggregate exposures and if EPA does not include them in the exposure assessment, it will not be able to accurately assess risks, including to potentially exposed or susceptible sub-populations. EPA has included exposure pathways under other environmental statutes administered by EPA in previous risk assessments, such as the example given above for the sulfuryl fluoride risk assessment that included fluoride exposures from drinking water.
- EPA states that exposures are "effectively managed" under other statutes, implying that these exposures do not present an unreasonable risk—but EPA did not provide any evidence to support this claim. Further, EPA did not evaluate the level of health risk related to these exposures as required by TSCA. Under TSCA, EPA must conduct a risk evaluation to determine if an unreasonable risk exists, without consideration of costs or other non-risk factors, including to potentially exposed or susceptible sub-populations. If an unreasonable risk exists, TSCA mandates that EPA make a rule to remove the unreasonable risk, [25] even if some of the exposures that contribute to the unreasonable risk are managed under other environmental statutes.

Hubal, E.C., Sheldon, L.S., Burke, J.M., McCurdy, T.R., Berry, M.R., Rigas, M.L., Zartarian, V.G. and Freeman, N.C., 2000. Children's exposure assessment: a review of factors influencing Children's exposure, and the data available to characterize and assess that exposure. Environmental health perspectives, 108(6), p.475.

[23] Abdallah MA-E, Harrad S. Dermal contact with furniture fabrics is a significant pathway of human exposure to brominated flame retardants. Environ Int. 2018 Sep 1;118:26–33.

[24] This statement is included in all 10 problem formulations. See, for example, US EPA (May 2018) Problem Formulation of the Risk Evaluation for Asbestos. Pg. 36

[25] 15 USC §2605

A rule for a chemical under another statute may not remove an unreasonable health risk because other statutes have different mandates for what EPA must consider when making a rule. For example, for carbon tetrachloride, EPA states that:

> "Carbon tetrachloride is a HAP [hazardous air pollutant]. EPA has issued a number of technology-based standards for source categories that emit carbon tetrachloride to ambient air and, as appropriate, has reviewed or is in the process of reviewing remaining risks. Because stationary source releases of carbon tetrachloride to ambient air are adequately assessed and any risks effectively managed when under the jurisdiction of the CAA [Clean Air Act], EPA does not expect to include emission pathways to ambient air from commercial and industrial stationary sources or associated inhalation exposure of the general population or terrestrial species in this TSCA evaluation."[26]

However, the statute for setting Maximum Achievable Control Technology (MACT) standards for hazardous air pollutants under the Clean Air Act does not require a risk evaluation, and the mandate is for the standard to achieve the reduction in emissions possible, considering technology, costs, and energy requirements—in contrast to TSCA, the CAA statute does *not* require that an unreasonable health risk be removed by the MACT emission standard.[27]

Indeed, a review of EPA's air toxics data reveals that every census tract in the U.S. has excess cancer risk of about 3.5 in a million due to carbon tetrachloride in the air— this is 3 times greater than what EPA typically considers an unreasonable cancer risk (1 in a million).[28]

Similarly for trichloroethylene (TCE), EPA states it does not plan to include inhalation of TCE from ambient air in the risk evaluation because risks are "effectively managed" under the Clean Air Act.[29] Yet EPA's air toxics data reveals cancer risks up to 19 in a million.[30]

This data contradicts EPA's unstated assertions that carbon tetrachloride and trichloroethylene in the air do not present an unreasonable risk to the general population.

Congress was aware of these other environmental statutes as the time of the passage of the 2016 amendments to TSCA and did not provide an exemption where other statutes addressed chemicals. On the contrary, the TSCA law directs EPA to conduct a risk evaluation that includes all exposures to determine if an unreasonable risk exists, and if it does, enact rules to remove the unreasonable risk. EPA must include chemical exposures from air, water, soil and other media for all the 10 chemicals, regardless of whether standards or rules exist under other environmental statutes.

1. **(d) EPA must assess risks to the general population for all 10 chemicals, and specifically to groups in proximity to conditions of use as a potentially exposed sub-population. Existing risks to the**

---

[26] US EPA (May 2018) Problem Formulation of the Risk Evaluation for Carbon Tetrachloride. pg. 48

[27] 42 USC §7412 (d)(2)-(3)

[28] US EPA (2011) National Air Toxics Assessment: 2011 NATA Assessment Results, Pollutant Specific Results: Carbon Tetrachloride. Available: https://www.epa.gov/national-air-toxics-assessment/2011-nata-assessment-results#pollutant

[29] US EPA (May 2018) Problem Formulation of the Risk Evaluation for Trichloroethylene. Pg. 54

[30] US EPA (2011) National Air Toxics Assessment: 2011 NATA Assessment Results, Pollutant Specific Results: Trichloroethylene. Available: https://www.epa.gov/national-air-toxics-assessment/2011-nata-assessment-results#pollutant

general population will be ignored for the following chemicals, as the problem formulations propose to exclude the general population entirely: PERC; asbestos; TCE; NMP; methylene chloride; carbon tetrachloride; 1,4 dioxane; and pigment violet 29.

Eight of the ten problem formulations wrongly propose to exclude the general population from any assessment of risk, including potentially exposed or susceptible sub-populations within the general population. As a result, EPA will not determine whether these eight chemicals pose unreasonable risks to pregnant women, infants, children, families living near current and former industrial sites, or any other sub-group within the general population who could potentially face health risks from chemical exposures.

For example, EPA notes a number of potentially exposed or susceptible sub-populations within the general population for PERC:[31]

- Other groups of individuals within the general population who may experience greater exposures due to their proximity to conditions of use identified in Section 2.2 that result in releases to the environment and subsequent exposures (e.g., individuals who live or work near manufacturing, processing, distribution or use sites)
- Perchloroethylene is lipophilic, and accumulates in fatty fluids and tissues in the human body. Subpopulations that may have higher body fat composition, and may be more highly exposed include pubescent and adult women, including women of child-bearing age. The EPA IRIS Assessment for perchloroethylene also identified the developing fetus as potentially exposed, as well as infants consuming breastmilk, particularly from mothers with occupational exposure to perchloroethylene or exposure due to proximity to industrial or commercial sources.
- Infants fed by formula may also experience increased perchloroethylene exposure if perchloroethylene is present in drinking water supplies.

Yet, EPA then states that it plans to ignore all these populations: "EPA does not expect to consider and analyze general population exposures in the risk evaluation for perchloroethylene. EPA has determined that the existing regulatory programs and associated analytical processes have addressed or are in the process of addressing potential risks of perchloroethylene that may be present in various media pathways (e.g., air, water, land) for the general population."[32] This is contrary to the TSCA mandate to consider risks to potentially exposed or susceptible sub-populations, and also not consistent with the current science on susceptibility. Fetuses, infants and children in the general population may not have the highest exposures to PERC or other chemicals, yet may actually be at greatest risk because their developing brains and bodies are the most vulnerable to a toxic chemical's effects.

To determine whether potentially exposed or susceptible sub-populations face unreasonable risks, EPA must assess risk to the general population for all 10 chemicals. People who live or work near former or current manufacturing, processing, distribution, use or disposal sites must be considered as facing greater exposures. It is well documented, including in EPA's own data and studies, that such communities are often low income and/ or people of color, exposed to a disproportionate share of

[31] US EPA (May 2018) Problem Formulation of the Risk Evaluation for Perchloroethylene. Pg. 47-8
[32] US EPA (May 2018) Problem Formulation of the Risk Evaluation for Perchloroethylene. Pg. 73

pollution, environmental hazards, social and economic stressors.[33,34] Multiple exposures to chemical and non-chemical stressors collectively increase the risk of harm, combined with synergistic effects with other health stressors in their daily lives such as limited access to quality health care.[35,36] In addition to the TSCA mandate to consider populations with greater exposure, the requirements of Executive Order 12898: Federal Actions To Address Environmental Justice in Minority Populations and Low-Income Populations also apply.[37] EPA should consult with the National Environmental Justice Advisory Council to ensure appropriate consideration of environmental justice populations in the TSCA risk evaluations.

1. (e) EPA should include realistic occupational exposure scenarios and not assume the use of exposure controls or compliance with mandatory or voluntary exposure standards.

For all the chemicals except pigment violet 29, EPA plans to assess occupational and occupational non-user exposures. For each chemical, EPA references the following exposure limits where they exist:

- Occupational Safety and Health Administration (OSHA) Permissible Exposure Limit (PEL) (regulatory)
- National Institute of Occupational Safety and Health (NIOSH) Recommended Exposure Limit (REL) (voluntary)
- American Conference of Government Industrial Hygienists (ACGIH) Threshold Limit Value (TLV) (voluntary)

For all the exposure assessments, EPA states that it plans to consider the influence of these existing exposure limits on occupational exposures.[38] It is not scientifically valid to assume that exposure limits represent the true distribution of worker exposures. First, EPA does not provide any data to support its assumption that existing exposure limits can be assumed to be the same as existing exposures. Exposure-monitoring data, across industries, large, small, and geographically diverse, and across all relevant job categories, are lacking. Second, compliance monitoring data, even when available, are not representative of worker exposure beyond the workplace where they are obtained. Even within a workplace, there is extremely high variability within and between worker exposures, and the results of compliance monitoring data are highly dependent on the number of samples taken.[39] For the purpose of

[33] US Environmental Protection Agency OAQPS, Institution (2015) Regulatory Impact Assessment of Final Revisions to the National Ambient Air Quality Standards for Ground Level Ozone EPA-452/R-15-007. Research Triangle Park, NC.

[34] Schulz, Amy J, Ments, GB, Sampson, Natalie, Ward, M, Anderson, R, deMajo, R, Isreal, BA, Lewis TC, Wilkins, D. (2016) Race and the distribution of social and physical environmental risk: a case example from the Detroit Metropolitan Area. DuBois Rev.

[35] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[36] Vesterinen HM, Morello-Frosch R, Sen S, Zeise L, Woodruff TJ. Cumulative effects of prenatal-exposure to exogenous chemicals and psychosocial stress on fetal growth: Systematic-review of the human and animal evidence. Meliker J, editor. PLoS One. 2017 Jul 12;12(7):e0176331.

[37] 59 FR 7629; February 16, 1994

[38] For example, from US EPA (May 2018) Problem Formulation of the Risk Evaluation for 1-Bromopropane. Pg. 64 "EPA will consider the influence of the recommended exposure limits on occupational exposures in the occupational exposure assessment."

[39] Rappaport SM. The rules of the game: An analysis of Osha's enforcement strategy. Am J Ind Med. 1984;6(4):291–303.

scientifically valid hazard and risk assessment, EPA needs to have data that describe the underlying distribution of exposures across the industries and occupations.

EPA does not provide any evidence to show that even OSHA PELs, which are enforceable, are routinely followed. Further, RELs and TLVs are voluntary and non-enforceable, so any adherence to such limits is subject to change in the future. Worker fatalities from methylene chloride[40] (which has an OSHA PEL) and complete disability from 1-bromopropane[41] (which has a TLV) demonstrate that exposure limits are doing little to protect workers from unreasonable risks.

PELs, RELs, TLVs and other administrative or engineering controls require people to give and get the right training and equipment, suppliers being transparent about the hazards in their products, and employers doing the right thing all the time (e.g., providing equipment, maintaining ventilation systems). We know this does not happen in real life, as engineering/ administrative controls may fail or not be adequate, such as the incorrect or improperly fitting personal protective equipment (PPE).

When evaluating occupational exposures, EPA needs to take into consideration all potential routes of exposure, and should not exclude exposure routes based on assumptions of PPE and/ or exposure controls in place. EPA should use monitoring data and relevant exposure models to estimate exposures, and not assume adherence to exposure limits. These controls are not guaranteed; assuming zero or limited exposure would be inappropriate and a failure to adequately ensure health protections, especially for potentially exposed and susceptible sub-populations as required by TSCA.

2. EPA should identify potential susceptible sub-populations based on established, scientifically supported extrinsic and intrinsic factors that increase vulnerability.

As shown below, EPA had identified known susceptibilities for certain chemicals in the Scoping documents, which are now removed with no explanation from the problem formulations.

Chemical	Susceptibilities identified in Scoping and removed from Problem Formulation
Asbestos	Age, pre-existing health conditions, genetic makeup/ genetic polymorphisms, co-exposure to other substances, early age at exposure, smoking, pre-existing respiratory conditions[42]
1-BP	Adult women of childbearing age and their offspring[43]
TCE	Life stage, gender-specific, genetic variation, race/ethnicity, preexisting health status, lifestyle factors and nutrition status[44]
Methylene Chloride	Life stage, gender-specific, genetic variation, preexisting health status, lifestyle factors, nutrition status, genetic polymorphisms[45]

[40] Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report. Fatal Exposure to Methylene Chloride Among Bathtub Refinishers — United States, 2000–2011. February 24, 2012.
[41] Urbina, Ian. As OSHA Emphasizes Safety, Long-Term Health Risks Fester. New York Times. March 30, 2013.
[42] US EPA (June 2017) Scope of the Risk Evaluation for Asbestos. Pg. 35
[43] US EPA (June 2017) Scope of the Risk Evaluation for 1-Bromopropane. Pg. 34
[44] US EPA (June 2017) Scope of the Risk Evaluation for Trichloroethylene. Pg. 38
[45] US EPA (June 2017) Scope of the Risk Evaluation for Methylene Chloride. Pg. 40

These factors are known susceptibilities for these chemicals. EPA's removal of the factors from the problem formulations is scientifically inappropriate and they should be considered in the risk evaluations.

Further, despite the TSCA requirements, EPA failed to identify susceptible sub-populations in the problem formulations for 1-BP, HBCD, asbestos, TCE, NMP, methylene chloride, 1,4-dioxane and pigment violet 29. EPA makes a general statement that it will evaluate available data to identify susceptible subpopulations.[46] This is not the scientifically appropriate approach, as the following are well-known factors that increase biologic sensitivity or reduce resilience to exposures.[47,48] Populations with these and other established factors should be considered a susceptible sub-population for each chemical, unless there is chemical-specific data showing otherwise.

Intrinsic/endogenous factors

- Genetic polymorphisms/genetics/genetic makeup
- Health status/nutritional status/disease status/pre-existing conditions
- Prenatal life stage
- Age

Extrinsic factors

- Multiple exposures/co-exposures
- Race/ethnicity
- Socioeconomic status (SES)

For example, the prenatal life stage can be the most sensitive to developmental and reproductive toxicants, and women of child-bearing age should be considered as a susceptible sub-population for any chemicals with such hazards. Yet, women of reproductive age are not identified as a potential susceptible sub-population for 1-BP, pigment violet 29, TCE, NMP, PERC, or HBCD, even though EPA will consider reproductive and developmental toxicity hazards for these chemicals.

EPA should apply a consistent and science-based approach to addressing susceptibility across the 10 chemicals by making the above and other relevant factors standard considerations for all 10 chemicals to identify susceptible sub-populations.

3. EPA should request additional test data for pigment violet 29 as the available data does not meet the TSCA requirement of adequate information to make a risk determination.

For pigment violet 29, EPA states that it plans to draw conclusions about hazards and exposures without further information or analysis: "EPA expects to be able to reach conclusions about particular conditions of use, hazards, or exposure pathways without further analysis and therefore plans to

[46] EPA states ""In developing the hazard assessment, EPA will evaluate available data to ascertain whether some human receptor groups may have greater susceptibility than the general population to the chemical's hazard(s)." See, for example, US EPA (May 2018) Problem Formulation for Cyclic Aliphatic Bromides Cluster (HBCD). Pg. 44

[47] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[48] National Research Council (2009) Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press.

conduct no further analysis on those conditions of use, hazards or exposure pathways in order to focus the Agency's resources on more extensive or quantitative analyses."[49]

TSCA statute[50] and regulation[51] requires adequate information to make a determination of whether or not a chemical poses an unreasonable risk. Regulation also requires the evaluation of "relevant" potential human and environmental hazards.[52]

To assess whether the data available on pigment violet 29 are 'adequate,' EPA should compare the completeness of the database on this chemical to existing lists of traits deemed important to assess for chemical safety. The health hazard dataset needed for EPA's Design for the Environment (DfE) program to conduct an alternatives assessment is such a data set. The necessary data is "to inform substitution to safer alternatives and reduce the likelihood of unintended consequences that might result if poorly understood alternatives were chosen."[53] The table below compares the empirical data available on pigment violet 29 with the requirements for a DfE human health hazard trait assessment.

DfE Hazard Trait	**Empirical Data Available for Pigment Violet 29?[54]**
Acute mammalian toxicity	Yes. In vivo oral, dermal and inhalation acute toxicity studies are available, though the inhalation studies are deemed to be unsuitable by ECHA.[55]
Respiratory sensitization	No
Skin sensitization	Yes, in vivo study
Eye irritation/ corrosivity	Yes, in vivo study
Skin irritation/ corrosivity	Yes, in vivo study
Carcinogenicity	No
Mutagenicity/ genotoxicity	Yes. In vitro gene mutation and mammalian cells genetic toxicity studies available.
Reproductive and developmental toxicity	Yes, screening study
Developmental neurotoxicity	No
Neurotoxicity	No
Repeated dose toxicity	No
Endocrine activity	No

Certain health hazards are specifically designated in TSCA statute, indicating that Congress expressly recognized these types of health effects as an unreasonable risk, and envisioned that EPA should assess them: "cancer/ carcinogenesis, mutagenesis/ gene mutation, teratogenesis, behavioral disorders, and

[49] US EPA (May 2018) Problem Formulation of the Risk Evaluation for Pigment Violet 29. Pg. 7
[50] 15 USC §2601 (b)(1)
[51] 40 CFR § 702.41 (b)
[52] 40 CFR § 702.41 (d)(3)
[53] EPA (2011) Design for the Environment Alternatives Assessment Criteria for Hazard Evaluation. Available: https://www.epa.gov/sites/production/files/2014-01/documents/aa_criteria_v2.pdf
[54] Information from: US EPA (May 2018) Problem Formulation of the Risk Evaluation for Pigment Violet 29. European Chemicals Agency (ECHA). (2017). Perylene-3, 4; 9, 10-tetracarboxydiimide. Helsinki, Finland. Available: https://echa.europa.eu/registration-dossier/-/registered-dossier/10330
[55] ECHA states: "Unsuitable test system, as the inhalation hazard test is insufficient for non-volatile substances." Available: https://echa.europa.eu/registration-dossier/-/registered-dossier/10330/7/3/3/?documentUUID=34aa4522-b714-47b0-9bee-af8052fff73d

birth defects."[56] EPA does not have empirical data on the carcinogenicity of pigment violet 29, nor on developmental neurotoxicity or endocrine activity, both of which are relevant to teratogenesis, behavioral disorders and birth defects.

Despite the lack of data on carcinogenicity, EPA states that: "However, negative genotoxicity results, SAR considerations and the expected negligible absorption and uptake of C.I. Pigment Violet 29, support EPA's conclusion that C.I. Pigment Violet 29 is unlikely to be a carcinogen."[57] This conclusion is not appropriate for two reasons.

First, absorption and uptake (or bioavailability) are exposure, not hazard considerations. Hazard traits are intrinsic properties of chemicals, while bioavailability relates to a chemical's exposure potential. Risk evaluations should assess hazard and exposure separately, then integrate the information to determine risks, as described in EPA's risk evaluation rule.[58] It is not appropriate for EPA to use "expected negligible absorption and uptake" to dismiss potential carcinogenicity— carcinogenicity hazard can only be demonstrated by data, as described below.

Second, according to the EPA Cancer Guidelines, the available data on pigment violet 29 are not adequate to support the conclusion that is "unlikely" to be a carcinogen. A determination of "Not Likely to Be Carcinogenic to Humans" requires **robust evidence** as follows:

> "This descriptor is appropriate when the available data are considered robust for deciding that there is no basis for human hazard concern. In some instances, there can be positive results in experimental animals when there is strong, consistent evidence that each mode of action in experimental animals does not operate in humans. In other cases, there can be convincing evidence in both humans and animals that the agent is not carcinogenic. The judgment may be based on data such as:
>
> - animal evidence that demonstrates lack of carcinogenic effect in both sexes in well-designed and well-conducted studies in at least two appropriate animal species (in the absence of other animal or human data suggesting a potential for cancer effects),
> - convincing and extensive experimental evidence showing that the only carcinogenic effects observed in animals are not relevant to humans,
> - convincing evidence that carcinogenic effects are not likely by a particular exposure route (see Section 2.3), or
> - convincing evidence that carcinogenic effects are not likely below a defined dose range.
>
> A descriptor of "not likely" applies only to the circumstances supported by the data. For example, an agent may be "Not Likely to Be Carcinogenic" by one route but not necessarily by another. In those cases that have positive animal experiment(s) but the results are judged to be not relevant to humans, the narrative discusses why the results are not relevant."[59]

Therefore, following the criteria established by the EPA, to determine that pigment violet 29 is not likely to be a carcinogen, supporting data from male and female animals of at least two species in well-

[56] 15 USC §2603 (b)(2)(A); 15 USC §2603 (e); 15 USC §2605 (b)(2)(D)
[57] US EPA (May 2018) Problem Formulation of the Risk Evaluation for Pigment Violet 29. Pg. 29
[58] 40 CFR §702.41
[59] U.S. Environmental Protection Agency (2005) Guidelines for Carcinogen Risk Assessment. Pg. 84-85. Available from: https://www.epa.gov/sites/production/files/2013-09/documents/cancer_guidelines_final_3-25-05.pdf

designed and conducted studies would be required. Negative genotoxic data and SAR considerations are not sufficient data to come to this final conclusion.

EPA's plan to complete the risk evaluation with the currently available data does not meet TSCA requirements, and EPA should request, at a minimum, testing data on pigment violet 29 carcinogenicity, developmental neurotoxicity, neurotoxicity, repeated dose toxicity and endocrine activity to ensure it has adequate information to complete the risk evaluation.

4. EPA should move forward with finalizing rules to limit uses of TCE and NMP immediately. EPA should not re-evaluate conditions of use of TCE and NMP already found to pose unreasonable risks.

EPA has completed risk evaluations for certain uses of TCE,[60] NMP[61] and methylene chloride[62] which were peer-reviewed and finalized. The conclusions of unreasonable risk in these existing evaluations are robust and meet TSCA standards. We commented in support of the science in the TCE risk evaluation in April 2017,[63] and EPA clearly believes that the science in these risk evaluations meets statutory and regulatory obligations, as the Agency states that it "intends to finalize the methylene chloride rulemaking proposed in January 2017," and that it will "…not re-evaluate the paint stripping uses of methylene chloride and will be relying on the previous assessment." [64]

Therefore, EPA's explanation for why it *will* re-evaluate the TCE and NMP uses already covered in the finalized risk assessments does not make sense: "EPA has concluded that the Agency's assessment of the potential risks from these widely used chemicals will be more robust if the potential risks from these conditions of use are evaluated by applying standards and guidance under amended TSCA."[65] The TCE, NMP and methylene chloride risk assessments were all completed using the same scientific standards and peer review processes. If the methylene chloride assessment meets TSCA standards and will be used to support rulemaking, [66] the same holds for the TCE and NMP assessments. EPA should not re-evaluate these TCE and NMP uses and instead proceed with rulemaking under the findings that these uses pose an unreasonable risk.

[60] US EPA (June 2014) TSCA Work Plan Chemical Risk Assessment Trichloroethylene: Degreasing, Spot Cleaning and Arts & Crafts Uses. Available: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemical-risk-assessment

[61] US EPA (March 2015) TSCA Work Plan Chemical Risk Assessment N-Methylpyrrolidone (NMP). Available: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemical-risk-assessment-n-0

[62] US EPA (Aug 2014) TSCA Work Plan Chemical Risk Assessment Methylene Chloride: Paint Stripping Use. Available: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemical-risk-assessment-methylene

[63] UCSF PRHE (April 2017) Comments from Academics, Scientists and Clinicians on Trichloroethylene (TCE): Regulation of Use in Vapor Degreasing Under TSCA Section 6(a). Available: https://www.regulations.gov/document?D=EPA-HQ-OPPT-2016-0387-0700

[64] US EPA (May 2018) EPA's Responses to Public Comments Received on the Scope Documents for the First Ten Chemicals for Risk Evaluation under TSCA. Pg. 15

[65] US EPA (May 2018) EPA's Responses to Public Comments Received on the Scope Documents for the First Ten Chemicals for Risk Evaluation under TSCA. Pg. 15

[66] US EPA (May 2018) EPA's Responses to Public Comments Received on the Scope Documents for the First Ten Chemicals for Risk Evaluation under TSCA. Pg. 15

Further, EPA states that it will re-evaluate the science on the TCE and NMP uses by applying the methods in the EPA document "Application of Systematic Review in TSCA Risk Evaluations." [67] This is not appropriate as (1) EPA is currently taking public comment on this document. It has not been peer-reviewed or finalized and (2) the document has major scientific and technical flaws which preclude its use for TSCA risk evaluations—see our comments on the systematic review document, attached as an appendix to these comments.

There is no scientific question that the uses of TCE, NMP and methylene chloride already evaluated by EPA pose unreasonable risks—indeed, the risk evaluations likely underestimate the magnitude of the unreasonable risk because not all exposure pathways are accounted for and risk values were not developed for non-cancer effects. EPA should quickly finalize the proposed rules to limit uses of TCE, NMP and methylene chloride to protect the public. Unless and until such uses are banned, the exposures from these uses continue and pose serious threats to public health—since January 2017, there have been at least 3 documented fatalities caused by methylene chloride.[68] Therefore, the new risk evaluations should use the exposure values from the previous assessments to consider the contributions of these uses to aggregate exposures.

[67] US EPA (May 2018) EPA's Responses to Public Comments Received on the Scope Documents for the First Ten Chemicals for Risk Evaluation under TSCA. Pg. 15

[68] Hopkins, J. "Reversing course, the EPA will regulate a deadly paint stripper." May 10, 2018. Center for Public Integrity. Available: https://www.publicintegrity.org/2018/05/10/21744/methylene-chloride-epa-regulation

Appendix
Comments from Academics, Scientists and Clinicians on: The Application of Systematic Review in TSCA Risk Evaluations

August 16, 2018

Comments from Academics, Scientists and Clinicians on: The Application of Systematic Review in TSCA Risk Evaluations.

Submitted online via *Regulations.gov* to docket EPA-HQ-OPPT-2018-0210

These comments are submitted on behalf of the undersigned academic, scientists, and clinicians. We declare collectively that we have no direct or indirect financial or fiduciary interest in any chemical under consideration in these risk evaluations. The co-signers' institutional affiliations are included for identification purposes only and do not necessarily imply any institutional endorsement or support, unless indicated otherwise.

We appreciate the opportunity to provide written comments on the Application of Systematic Review in TSCA Risk Evaluations,[a] pursuant to the Toxic Substances Control Act (TSCA), as amended by the Frank R. Lautenberg Chemical Safety of the 21st Century Act (Lautenberg TSCA). TSCA requires that EPA make decisions about chemical risks based on the "best available science" and the "weight of the scientific evidence"[b] which EPA defined in regulation as "…a systematic review method, applied in a manner suited to the nature of the evidence or decision, that uses a pre-established protocol to comprehensively, objectively, transparently, and consistently identify and evaluate each stream of evidence, including strengths, limitations, and relevance of each study and to integrate evidence as necessary and appropriate based upon strengths, limitations, and relevance.''[c]

Systematic review methods originated more than 40 years ago in psychology. The methodology was soon adapted to evaluating the effectiveness of clinical interventions in medicine and related disciplines in response to empirical evidence demonstrating the need to apply scientific principles not only to primary research, but also to research synthesis methods that inform decision-making in healthcare (1-3). Almost a decade ago, these empirically-proven methods for research synthesis were adapted to environmental health (4, 5). To date, science-based methods for systematic review in environmental health have been demonstrated in case studies in the peer-reviewed literature (6-13), and adopted by the National Toxicology Program (14) and the U.S. EPA's Integrated Risk Information System (IRIS) program (15).

EPA's systematic review framework under TSCA establishes EPA's "rules" for assembling and interpreting the scientific evidence on chemicals in commerce. These "rules" will determine, whether explicitly, implicitly, and/or by default, *what* evidence EPA will consider, and *how* it will evaluate that evidence when it is making decisions about potentially hazardous chemicals in commerce. Exposure to industrial, commercial, and consumer product chemicals is ubiquitous from the time of conception until death. As such, EPA's rules for gathering and interpreting the science that evaluates the relationship between these exposures and adverse health effects are of profound importance to the general public, and will have even greater impact on the potentially exposed or susceptible sub-populations Congress explicitly mandated EPA to protect: pregnant women, children, individuals with underlying health conditions, workers, and those with greater exposure and/or greater vulnerability to chemical toxicity and exposure.

[a] 83 FR 26998, June 11, 2018
[b] 15 USC §2625 (h)-(i)
[c] 40 CFR 704.33

With so much at stake, we are deeply concerned by EPA's ad hoc and incomplete TSCA systematic review framework, which is inconsistent with current, established, best available empirical methods for systematic review. Moreover, as we detail below, the application of EPA's TSCA framework would likely result in the exclusion of quality research from EPA's decision-making. Accordingly, the TSCA systematic review method does not meet the mandate of the law to use the "best available science." [d]

Based on the most current empirically demonstrated principles of systematic review methods, we provide EPA with concrete recommendations and approaches to correct its methodology and inform timely science-based decision-making to achieve the Agency's mission of protecting the public from harmful chemicals.

Our comments address the following six main points:

1. **EPA's TSCA systematic review framework is ad hoc, incomplete, and does not follow established methods for systematic review that are based on the best available science.**

   We recommend: EPA should implement a systematic review method that is compatible with empirically based existing methods and aligns with the Institute of Medicine's[e] definition of a systematic review, including but not limited to, using explicit and pre-specified scientific methods for every step of the review. EPA should consider methods demonstrated for use in environmental health, and which have been endorsed and utilized by the National Academy of Sciences, i.e., the National Toxicology's Office of Heath Assessment and Translation systematic review method, and the Navigation Guide Systematic Review Method. EPA's TSCA systematic review framework should be peer-reviewed by qualified external experts in the field.

2. **EPA's TSCA systematic review framework utilizes a quantitative scoring method that is incompatible with the best available science in fundamental ways:**

   a. **Quantitative scores for assessing the quality of an individual study are arbitrary and not science-based; the Cochrane Collaboration and National Academy of Sciences recommend against such scoring methods.**
   b. **EPA's scoring method wrongly conflates how well a study is reported with how well the underlying research was conducted; and**
   c. **EPA's scoring method excludes research based on one single reporting or methodological limitation.**

   We recommend: EPA should not use a quantitative scoring method to assess quality in individual studies; it should not conflate study reporting with study quality; and it should not exclude otherwise quality research based on a single reporting or methodological limitation. Rather EPA should employ a scientifically valid method to assess risk of bias of individual studies.

3. **EPA's TSCA systematic review framework does not consider financial conflicts of interest as a potential source of bias in research.**

[d] 15 USC §2625 (h)

[e] The Institute of Medicine is now the National Academy of Medicine.

We recommend: EPA should assess study and author funding source as a risk of bias domain for individual studies.

4. **The literature review step of EPA's TSCA systematic review framework incorporates select best practices, but also falls short of, or is unclear about, many other best practices for conducting a systematic and transparent literature review.**

   We recommend: EPA should make its framework for conducting a literature review congruent with all of the Institute of Medicine's best practices and explicitly include rules for when the list of relevant studies will be considered final.

5. **EPA's TSCA systematic review framework correctly recognizes that mechanistic data are not required for a hazard assessment, but EPA is not clear that these data, if available, can only be used to increase, and not to decrease, confidence in a body of evidence.**

   We recommend: EPA should be explicit that mechanistic data can only be used to upgrade a hazard classification, or increase the confidence of a finding made based on evaluation of animal and human data, and that these data will not be used to decrease confidence in a body of evidence.

6. **EPA's TSCA systematic review framework is not independent of the regulatory end user of the review.**

   We recommend: EPA's TSCA systematic reviews should be produced independently of the regulatory end user of the review.

We are appreciative of the opportunity to provide public input. Please do not hesitate to contact us with any questions regarding these comments.

Sincerely,

Veena Singla, PhD
Associate Director, Science and Policy, Program on Reproductive Health and the Environment
University of California, San Francisco

Patrice Sutton, MPH
Research Scientist, Program on Reproductive Health and the Environment
University of California, San Francisco

Tracey Woodruff, PhD, MPH
Director, Program on Reproductive Health and the Environment
University of California, San Francisco

Juleen Lam, PhD, MHS, MS
Assistant Professor, Department of Health Sciences
California State University, East Bay

Patricia D. Koman, PhD, MPP
President and Senior Health Scientist
Green Barn Research Associates*
Ann Arbor, Michigan

Lisa Bero, PhD
Chair of Medicines Use and Health Outcomes, Charles Perkins Centre
The University of Sydney

Liz Borkowski, MPH
Senior Research Scientist, Milken Institute School of Public Health
George Washington University

Sheila Brear, BDS
Associate Dean, Academic Affairs, School of Dentistry
University of California, San Francisco

Adelita G. Cantu, PhD, RN
Associate Professor
Alliance of Nurses for Healthy Environments

Courtney Carignan, PhD
Assistant Professor
Michigan State University

Daniel M. Fox, PhD
President Emeritus
Milbank Memorial Fund

Danielle Fries, MPH
Science Associate, Program on Reproductive Health and the Environment
University of California, San Francisco

Mary Gant, MS
Retired Policy Analyst
National Institute of Environmental Health Sciences

Steven G. Gilbert, PhD, DABT
Affiliate Professor
University of Washington

Robert M. Gould, MD
Associate Adjunct Professor, Department of Obstetrics, Gynecology and Reproductive
University of California, San Francisco
Past-President, Physicians for Social Responsibility

Maeve Howett, PhD, APRN, CPNP, IBCLC, CNE
Clinical Professor and Assistant Dean

University of Massachusetts Amherst

Jyotsna Jagai, MS, MPH, PhD
Research Assistant Professor, School of Public Health
University of Illinois at Chicago

Paula I. Johnson, PhD, MPH
Research Scientist, Safe Cosmetics Program
California Department of Public Health

Jean-Marie Kauth, PhD, MPH
Professor
Benedictine University

Carol Kwiatkowski, PhD
Executive Director
The Endocrine Disruption Exchange*

Joseph Laakso, PhD
Director, Science Policy
Endocrine Society*

Gail Lee, RD, REHS Hem
Sustainability Director
University of California, San Francisco

Michael J. Martin, MD, MPH, MBA
Associate Clinical Professor
University of California, San Francisco

Rachel Morello-Frosch, PhD, MPH
Professor, School of Public Health and Department of Environmental Science, Policy and Management
University of California, Berkeley

Katherine Pelch, PhD
Senior Scientist
The Endocrine Disruption Exchange

Janet Pelrman, MD, MPH
Physician
Stanford Children's Hospital

Jeanne Rizzo, RN
President & CEO
Breast Cancer Prevention Partners

Ted Schettler, MD, MPH
Science Director

Science and Environmental Health Network

Rachel M. Shaffer, MPH
PhD Candidate, School of Public Health
University of Washington, Seattle

Laura N. Vandenberg, PhD
Associate Professor
University of Massachusetts, Amherst

Ellen M. Wells, PhD
Assistant Professor of Environmental & Occupational Health
Purdue University School of Health Sciences

Nsedu Obot Witherspoon, MPH
Executive Director
Children's Environmental Health Network

Marya Zlatnik, MD, MMS
Professor, Department of Obstetrics, Gynecology & Reproductive Sciences
University of California, San Francisco

*indicates organizational support

DETAILED COMMENTS

1. EPA's TSCA systematic review framework is ad hoc, incomplete, and does not follow established methods for systematic review that are based on the best available science.

The best available scientific method for a systematic review (SR) specifies that all components of a review be established in a publically available protocol written *prior* to conducting the review to minimize bias and to ensure transparency in decision-making. For example, the Institute of Medicine defines a systematic review as a "scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies" (emphasis added) (16)(p.1). A fatal flaw in EPA's SR framework is that it lacks essential SR elements, including but not limited to: (1) a protocol for executing a SR developed *prior* to conducting the SR; (2) an explicit method for evaluating the overall body of each evidence stream, i.e., animal, human, etc.; and (3) an explicit method for integrating two or more streams of evidence, including defined criteria for the type and level of evidence needed for a decision by EPA.

Notably, EPA's TSCA SR Framework presents a diagram of a complete SR framework in Figure 3-1 (page 15) and states in footnote 4 on that page that the:

> Diagram depicts systematic review process to guide the first ten TSCA risk evaluations. It is anticipated that the same basic process will be used to guide future risk evaluations with some potential refinements reflecting efficiencies and other adjustments adopted as EPA/OPPT gains experience in implementing systematic review methods and/or approaches to support risk evaluations within statutory deadlines (e.g., aspects of protocol development would be better defined prior to starting scoping/problem formulation).

However, EPA's TSCA SR Framework then proceeds to describe an ad hoc and highly flawed method limited to only the data collection and, to a limited extent, the data evaluation components of a SR. Specifically, Figure S-1 below, excerpted from the National Academy of Sciences 2014 review of the EPA IRIS program's systematic review method (17), presents all of the components of a science-based SR. The red box indicates the parts of a SR method that EPA has included in its proposed framework.

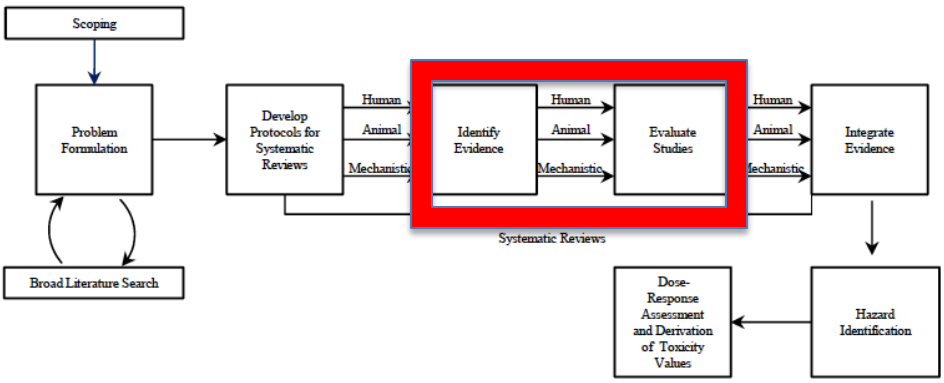


FIGURE S-1 Systematic review in the context of the IRIS process. The committee views public input and peer review as integral parts of the IRIS process, although they are not specifically noted in the figure.

EPA's piecemeal approach is not only in direct contradiction with the best available scientific methods for SR, but also incompatible with the regulatory definition of[f] "weight of evidence" in the risk evaluation rule, which specifies a complete method spelled out in a protocol developed *before* conducting the review. Therefore, the TSCA systematic review method violates both TSCA statute and regulation. [g]

EPA explicitly states that it is proceeding with its first ten risk assessments in the absence of a pre-defined protocol and a complete method for systematic review. Specifically, EPA's SR Framework states:

> (p. 9) … the purpose of the document is internal guidance that … sets out general principles to guide EPA's application of systematic review in the risk evaluation process for the first ten chemicals … **EPA had limited ability to develop a protocol document detailing the systematic review approaches and/or methods prior to the initiation of the risk evaluation process for the first ten chemical substances**. **For these reasons, the protocol development is staged in phases while conducting the assessment work" (emphasis added).** Additional details on the approach for the evidence synthesis and integration will be included with the publication of the draft TSCA risk evaluations.

In effect, EPA is saying it does not have time to comply with its regulatory requirement to conduct a science-based systematic review, and will not actually develop its protocol until it completes the first ten systematic reviews.

First, this approach is in clear violation with scientifically-validated approaches to conducting systematic reviews. In its review of the EPA's Integrated Risk Information System (IRIS) program's proposed SR methods, the National Academy of Sciences specified that, "Completing the literature search as part of

[f] EPA's risk evaluation rule (40 CFR 704.33) states: ''Weight of the scientific evidence means a systematic review method, applied in a manner suited to the nature of the evidence or decision, that uses a pre-established protocol to comprehensively, objectively, transparently, and consistently identify and evaluate each stream of evidence, including strengths, limitations, and relevance of each study and to integrate evidence as necessary and appropriate based upon strengths, limitations, and relevance.''

[g] 15 USC §2625 (h)-(i) and 40 CFR 704.33

protocol development is inconsistent with current best practices for systematic review, and the IRIS program is encouraged to complete the public-comment process and finalize the protocol before initiating the systematic review" (15)(Pg. 8). In the case of TSCA risk assessments, EPA is not only completing the literature search as part of protocol development, it is completing the entire systematic review in the absence of a protocol and complete method. It is blatantly biased to write the rules of evidence assembly and interpretation at the same time one is applying the rules, and as such, this method cannot be validly referred to as a science-based systematic review.

Second, a lack of time is not a credible rationale for EPA's failure to conduct a science-based systematic review for the first ten TSCA chemicals. There are multiple well-developed, science-based, peer-reviewed and validated methods for conducting systematic reviews in environmental health that EPA could readily apply, including the SR method and handbook developed by the Office of Health Assessment and Translation at the National Toxicology Program (14), and the Navigation Guide Systematic Review Method, which has been demonstrated in six case studies (6-13). The National Academy of Sciences cited both of these SR methods as exemplary of the type of methods EPA should use in hazard and risk assessment (17, 18). Further, the National Academy of Sciences utilized both methods in its 2017 assessment of the potential health impacts of endocrine active environmental chemicals (19). Specifically, in its 2017 review the National Academy of Sciences found:

> The two approaches [OHAT and Navigation Guide] are very similar … and they are based on the same established methodology for the conduct of systematic review and evidence assessment (e.g., Cochrane Collaboration, AHRQ Evidence-based Practice Center Program, and GRADE). Both the OHAT and Navigation Guide methods include the key steps recommended by a previous National Academies committee (NRC 2014) for problem formulation, protocol development, specifying a study question, developing PECO statement, identifying and selecting the evidence, evaluating the evidence, and integrating the evidence" (19)(page 119).

Protocols developed for applying the Navigation Guide and the OHAT method have been published and can serve as a template to further expedite EPA's TSCA reviews.[h]

Furthermore, the language of EPA's systematic review framework is confusing, contradictory, and poorly and incorrectly referenced with little science or policy foundation. This suggests the authors of EPA's TSCA Systematic Review Framework lack sufficient understanding of the scientific process integral to this work. A particularly egregious example is EPA's stated understanding of EPA's TSCA statutory science standards:

> (Pg. 26) EPA/OPPT is required by TSCA to use the weight of the scientific evidence in TSCA risk evaluations. Application of weight of evidence analysis is an integrative and interpretive process that considers both data/information in favor (e.g., positive study) or against (e.g., negative study) a given hypothesis within the context of the assessment question(s) being evaluated in the risk evaluation.

This directly contradicts EPA's own published rule which defines what a systematic review is (see

[h] All Navigation Guide systematic review protocols can be found at: https://prhe.ucsf.edu/navigation-guide The National Toxicology Program's protocol for its systematic review to evaluate the evidence for an association between exposure to PFOA or PFOS and immunotoxicity or immune-related health effects is at: https://ntp.niehs.nih.gov/ntp/ohat/pfoa_pfos/protocol_201506_508.pdf

footnote “e”, above) and such an understanding completely subverts the purpose of a systematic review which is to explicitly avoid a simplistic analysis that would led to erroneous conclusions along the lines of stating that, for instance, “five studies are in favor (positive) and ten are against (negative) and therefore the weight is … ”

Another bewildering statement by EPA concerns its highly quantitative scoring method, which is the main topic of its systematic review framework (see comment #2, below). EPA adds a caveat to the scoring method that says quantitative scoring is actually a qualitative method, and further: “The [scoring] system is not intended to imply precision and/or accuracy of the scoring results” (Pg. 35).

The ad hoc and incomplete nature of EPA’s systematic review framework is incompatible in many additional fundamental ways, described further in detail below, with science based methods of systematic review developed, endorsed, and/or advanced by the: National Academy of Sciences (17-19); the Institute of Medicine (16); the National Toxicology Program (14); the Cochrane Collaboration (20); the Grading of Recommendations Assessment, Development and Evaluation (GRADE) method (21, 22); the international scientific collaboration that developed a framework for the “systematic review and integrated assessment” (SYRINA) of endocrine disrupting chemicals (23); the SYRCLE systematic review method for animal studies (24); the Campbell Collaboration’s methods (25); and the Navigation Guide systematic review method developed by a collaboration of scientists led by the University of California San Francisco (4). Most of these organizations also pre-publish their protocols either online (i.e., the National Toxicology Program) or in PROSPERO[i] (i.e., UCSF).

We recommend: EPA should implement a systematic review method that is compatible with empirically based existing methods and aligns with the Institute of Medicine’s definition of a systematic review, including, but not limited to, using explicit and pre-specified scientific methods for every step of the review. EPA should consider methods demonstrated for use in environmental health, and which have been endorsed and utilized by the National Academy of Sciences, i.e., the National Toxicology’s Office of Heath Assessment and Translation systematic review method, and the Navigation Guide Systematic Review Method. EPA’s TSCA systematic review framework should be peer-reviewed by qualified external experts in the field.

[i] PROSPERO International prospective register of systematic reviews https://www.crd.york.ac.uk/prospero/

2. EPA's TSCA systematic review framework utilizes a quantitative scoring method that is incompatible with the best available science in fundamental ways:

- **a. Quantitative scores for assessing the quality of an individual study are arbitrary and not science-based; the Cochrane Collaboration and National Academy of Sciences recommend against such scoring methods.**
- **b. EPA's scoring method wrongly conflates how well a study is reported with how well the underlying research was conducted; and**
- **c. EPA's scoring method excludes research based on one single reporting or methodological limitation.**

A detailed explanation of each of these scientific shortcomings is provided below.

(a) Quantitative scores for assessing the quality of an individual study are arbitrary and not science-based.

EPA's SR framework employs a quantitative scoring method to assess the quality of individual studies, assigning, based on its "professional judgment", various weights for quality domains and then summing up the quantitative scores to decide whether a study is of "high", "medium", or "low" quality as follows:[j]

> (Pg. 33) A numerical scoring method is used to convert the confidence level for each metric into the overall quality level for the data/information source. The overall study score is equated to an overall quality level (*High*, *Medium*, or *Low*) using the level definitions and scoring scale shown in Table A-1. The scoring scale was obtained by calculating the difference between the highest possible score of 3 and the lowest possible score of 1 (i.e., 3-1= 2) and dividing into three equal parts (2 ÷ 3 = 0.67). This results in a range of approximately 0.7 for each overall data quality level, which was used to estimate the transition points (cut-off values) in the scale between *High* and *Medium* scores, and *Medium* and *Low* scores. These transition points between the ranges of 1 and 3 were calculated as follows: Cut-off values between *High* and *Medium*: 1 + 0.67= 1.67, rounded up to 1.7 (scores lower than 1.7 will be assigned an overall quality level of *High*) Cut-off values between *Medium* and *Low*: 1.67 + 0.67= 2.34, rounded up to 2.3 (scores between 1.7 and lower than 2.3 will be assigned an overall quality level of *Medium)*

<u>This overall scoring method is applied to all streams of evidence, and our comments reflect our objection to EPA's applying scoring to any and all streams of evidence</u>.[k]

Illustrative of the scoring method, in Appendix H "Data Quality Criteria for Epidemiologic Studies," (page

[j] See Appendix A for a more detailed description of the scoring method; how the method will be applied specifically to various streams of evidence, i.e., occupational exposure and release data; animal and in vitro data; epidemiologic studies; etc., is described in subsequent Appendices B-H.

[k] EPA's framework applies quantitative scoring to all types of data; EPA/OPPT "is not applying weighting factors to the general population, consumer, and environmental exposure data types. In practice, it is equivalent to assigning a weighting factor of 1, which statistically assumes that each metric carries an equal amount of weight." (Pg. 96).

225) EPA presents how scoring is further applied to human studies, explaining:

> The critical metrics within each domain are those that cover the most important aspects of the domain and are those that more directly evaluate the role of confounding and bias. After pilot testing the evaluation tool, EPA recognized that more attention (or weight) should be given to studies that measure exposure and disease accurately and allow for the consideration of potential confounding factors. Therefore, metrics deemed as critical metrics are those that identify the major biases associated with the domain, evaluate the measurement of exposure and disease, and/or address any potential confounding. … EPA/OPPT assigned a weighting factor that is twice the value of the other metrics within the same domain to each critical metric. Remaining metrics are assigned a weighting factor of 0.5 times the weighting factor assigned to the critical metric(s) in the domain. The sum of the weighting factors for each domain equals one.

There is no scientific evidence to support EPA's selection of these "critical metrics" as being more important that other metrics, i.e., why within the "study participation" domain "selection" and "attrition" are more important than "comparison group"; and there are no data supporting EPA's choice of particular numbers for weighting these 'critical metrics' (i.e., some metrics are "twice" as important as the other metrics).

Overall, there is no scientific justification for EPA to assign these or any other quantitative scoring measures for assessing the quality of an individual study. The implicit assumption in quantitative scoring methods is that we know empirically how much each risk of bias domain contributes to study quality, and that these domains are independent of each other. This is not a scientifically supportable underlying assumption. Research has documented that scoring methods have, at best, unknown validity, may contain invalid items, and that results of a quality score are not scientifically meaningful or predictive of the quality of studies (26-28). An examination of the application of quality scores in meta-analysis found that quality-score weighting produced biased effect estimates, with the authors explaining that quality is not a singular dimension that is additive, but that it is possibly non-additive and non-linear (29).

Aggregating across quality criteria to produce a single score is recognized by preeminent systematic review methodologists as problematic and unreliable because the weights assigned are arbitrary and focus on the quality of reporting rather than the design and conduct of the research (21, 30). Scoring is not utilized by empirically based systematic review methodologies, such as the Cochrane Collaboration or GRADE (21, 31). As stated by the Institute of Medicine, "… systematic review teams have moved away from scoring systems to assess the quality of individual studies toward a focus on the components of quality and risk of bias" (16).

The Cochrane Collaboration, founded in 1993, is an international non-profit and independent organization that produces and disseminates systematic reviews of healthcare interventions and is a key locus of the world's most authoritative expertise on systematic review methods. Cochrane's methodology states: "The current standard in evaluation of clinical research calls for reporting each component of the assessment tool separately **and not calculating an overall numeric score** **(emphasis added**)"(31).

The National Academy of Sciences in its review of the EPA's IRIS program's method for SR, strongly supported a methodology that did not incorporate quantitative scoring, stating:

> … Cochrane discourages using a numerical scale because calculating a score involves choosing a weighting for the subcomponents, and such scaling generally is nearly impossible to justify (Juni et al. 1999). Furthermore, a study might be well designed to eliminate bias, but because the study failed to report details in the publication under review, it will receive a low score. Most scoring systems mix criteria that assess risk of bias and reporting. However, there is no empirical basis for weighting the different criteria in the scores. Reliability and validity of the scores often are not measured. Furthermore, quality scores have been shown to be invalid for assessing risk of bias in clinical research (Juni et al. 1999). The current standard in evaluation of clinical research calls for reporting each component of the assessment tool separately and not calculating an overall numeric score (Higgins and Green 2008) (17)(Pg. 69).

b) EPA's scoring method wrongly conflates how well a study is reported with how well the underlying research was conducted.

Study reporting addresses how well research findings are written up, i.e., whether there is a complete and transparent description of what was planned, what was done, what was found, and what the results mean. Guidelines and checklists for authors have been developed to help ensure all information pertinent to assessing the quality and meaning of research is included in the report. The "Strengthening of Reporting of Observational Studies in Epidemiology" or "STROBE" Initiative is an example of a checklist of items that should be included in articles reporting such research.[l]

EPA's SR Framework uses reporting measures in its scoring of the quality of human studies, including incorporating reporting guidelines into the reasons for scoring studies "low quality" (Metrics 1 and 15) or "unacceptable for use" (Metrics 2, 3, 4, 6, 7). EPA's SR Framework acknowledges that reporting is not the same as an underlying flaw in study methodology (Pg. 31), but then proceeds to ignore this distinction by using reporting as a measure of the quality of the underlying research. EPA's SR Framework not only does not "untangle" reporting from quality, it specifically conflates the two by using metrics in the STROBE reporting guidelines to score individual studies. The authors of the STROBE guidelines specifically note the guidelines are not a measure of the quality of the underlying research, stating:

> The STROBE Statement is a checklist of items that should be addressed in articles reporting on the 3 main study designs of analytical epidemiology: cohort, case control, and cross-sectional studies. The intention is solely to provide guidance on how to report observational research well; these recommendations are not prescriptions for designing or conducting studies. Also, while clarity of reporting is a prerequisite to evaluation, the checklist is not an instrument to evaluate the quality of observational research (emphasis added). … Our intention is to explain how to report research well, not how research should be done. We offer a detailed explanation for each checklist item. Each explanation is preceded by an example of what we consider transparent reporting. This does not mean that the study from which the example was taken was uniformly well reported or well done; nor does it mean that its findings were reliable, in the sense that they were later confirmed by others: it only means that this particular item was well reported in that study."(32)

How completely and clearly a study is reported is not a scientifically valid measure of the quality of the

[l] See Strobe statement at: https://www.strobe-statement.org/index.php?id=strobe-aims

underlying research (20, 21, 33, 34). As GRADE methodologists have succinctly stated, "… just because a safeguard against bias is not reported does not mean it was neglected"(21). Moreover, including many reporting items that are irrelevant to bias in a quality scoring rule (e.g., an indicator of whether power calculations were reported), will disproportionately reduce some of the resulting scores (29).

The Cochrane Collaboration Handbook for conducting a SR clearly distinguishes reporting and bias, the latter which is defined as "a systematic error, or deviation from the truth, in results or inferences" (20). The Cochrane Manual for conducting systematic reviews is explicit about not conflating reporting with bias, stating:

> Bias may be distinguished from **quality**. The phrase 'assessment of methodological quality' has been used extensively in the context of systematic review methods to refer to the critical appraisal of included studies. The term suggests an investigation of the extent to which study authors conducted their research to the highest possible standards. This *Handbook* draws a distinction between assessment of methodological quality and assessment of risk of bias, and recommends a focus on the latter. The reasons for this distinction include:
>
> 1. The key consideration in a Cochrane review is the extent to which results of included studies should be *believed*. Assessing risk of bias targets this question squarely.
> 2. A study may be performed to the highest possible standards yet still have an important risk of bias. For example, in many situations it is impractical or impossible to blind participants or study personnel to intervention group. It is inappropriately judgemental to describe all such studies as of 'low quality', but that does not mean they are free of bias resulting from knowledge of intervention status.
> 3. Some markers of quality in medical research, such as obtaining ethical approval, performing a sample size calculation and reporting a study in line with the CONSORT Statement (Moher 2001d), are unlikely to have direct implications for risk of bias.
> 4. An emphasis on risk of bias overcomes ambiguity between the quality of reporting and the quality of the underlying research (although does not overcome the problem of having to rely on reports to assess the underlying research).

Importantly, in the application of EPA's SR Framework, studies can be scored as "low quality," and even excluded from EPA's review, based solely on a deficiency in reporting, irrespective of the quality of the underlying research. Research documents that important information is often missing or unclear in published research (35), as word limits, styles, and other specifications are highly variable, and non-standardized among peer-reviewed journals. As such, efforts to improve reporting are focused on uptake of reporting guidelines by journal editors and researchers (32, 36, 37). Improving reporting is needed in academic research, but as stated by the developers of the STROBE guidelines, "We want to provide guidance on how to report observational research well. … the checklist is not an instrument to evaluate the quality of observational research."

Given the historical and present-day deficiencies in how studies are reported in the peer-reviewed literature, and because EPA's scoring system rates as 'unacceptable for use' any human study that does not report even one of five reporting metrics, EPA's proposal could reasonably be expected to lead to the exclusion from EPA's consideration much of the existing body of knowledge on the impact of

environmental chemicals on human health, and is inconsistent with TSCA mandates to use the "best available science" and "reasonably available information." [m] Applying flawed exclusion criteria that directly contradicts widely accepted empirically based SR methodological approaches will almost certainly result in flawed conclusions and threaten the protection of the public's health.

(c) EPA's scoring method excludes research based on one single reporting or methodological limitation.

In the "fatal flaw" component of EPA's SR Framework's scoring system, for each type of evidence stream, i.e., epidemiologic, animal, *in vitro*, etc., EPA created an arbitrary list of metrics that make studies "unacceptable for use in the hazard assessment," stating:

> EPA/OPPT plans to use data with an overall quality level of *High, Medium*, or *Low* confidence to quantitatively or qualitatively support the risk evaluations, but does not plan to use data rated as *Unacceptable*. Studies with any single metric scored as 4 will be automatically assigned an overall quality score of *Unacceptable* and further evaluation of the remaining metrics is not necessary (emphasis added). An *Unacceptable* score means that serious flaws are noted in the domain metric that consequently make the data unusable (or invalid) (Pg. 227).

There is no empirical basis for EPA's selected list of fatal flaws.

Illustrative of this "fatal flaw" aspect of EPA' scoring system, for human epidemiologic studies (See Section H.5, Table H-8 (page 231), EPA lists six domains of study quality, i.e., study participation; exposure characterization; outcome assessment; potential confounding/variable control; analysis**;** and other considerations for biomarker selection and measurement, and 19 metrics to assess the six domains. A study that has even one of the 19 "serious flaws" metrics is considered to be "unacceptable for use."

The underlying assumptions of EPA's "serious flaws" metrics are not science-based because:

- **EPA's list of "serious flaws" are not all equal indicators of study quality**:
  For example, among human observational studies, any one of the list of 19 metrics can eliminate a study from consideration as EPA considers all of these "flaws" to be of equal import; as described in detail above, such weighting is arbitrary and not a science-based method.

- **EPA's list of "serious flaws" are not all related to real flaws in the underlying research**:

  - **Reporting** guidelines are wrongly equated with "serious flaws" in study quality.
    For example, in scoring the quality of human studies, 5 of 19 "serious flaw" metrics (Table H-8) are STROBE reporting guidelines (STROBE checklist items # 6,7,8,13,15). A study would be scored as "unacceptable for use" by EPA based on any one of these STROBE reporting guidelines. As described above in comment #2a, the STROBE guideline developers explicitly state this is neither the intended nor a scientifically valid use of these guidelines. (32)

[m] 15 USC §2625(h) and (k)

- **Analysis**[n] is equated with a "serious flaw" in study quality, but statistical power[o] alone is not a valid measure of study quality. For example, EPA's framework excludes human studies that do not meet EPA's criteria for "high" in the analysis domain. EPA does not state how it will calculate whether a study is "adequately" powered. According to EPA's framework, to be included in an EPA review, a study must meet the "high" criteria in EPA's "Metric 13, Statistical power (sensitivity, reporting bias)" as presented in the box below. Studies that are not "high" quality for this metric would be designated as "unacceptable for use" by EPA:

EPA Metric 13. Excerpted from Table H-9 (page 243)

Metric 13. Statistical power (sensitivity, reporting bias) **Instructions: To meet criteria for confidence ratings for metrics where 'AND' is included, studies must address both of the conditions where "AND" is stipulated. To meet criteria for confidence ratings for metrics where 'OR' is included studies must address at least one of the conditions stipulated.**	
High (score = 1)	***For cohort and cross-sectional studies:*** The number of participants are adequate to detect an effect in the exposed population and/or subgroups of the total population. **OR** The paper reported statistical power high enough (≥ 80%) to detect an effect in the exposure population and/or subgroups of the total population. ***For case-control studies:*** The number of cases and controls are adequate to detect an effect in the exposed population and/or subgroups of the total population. **OR** The paper reported statistical power was high (≥ 80%) to detect an effect in the exposure population and/or subgroups of the total population.
Medium (score = 2)	• Do not select for this metric.
Low (score = 3)	• Do not select for this metric.
Unacceptable (score = 4)	• ***For cohort and cross-sectional studies:*** The number of participants are inadequate to detect an effect in the exposed population and/or subgroups of the total population. • ***For case-control studies:*** The number of cases and controls are inadequate to detect an effect in the exposed population and/or subgroups of the total population.

[n] See Table H-8 "Serious Flaws that Would Make Epidemiological Studies Unacceptable for Use in the Hazard Assessment" under the "analysis domain" "statistical power/sensitivity" metric (page 233) " in conjunction with Table H-9 "Evaluation Criteria for Epidemiologic Studies, Metric 13 "statistical power (sensitivity, reporting bias) (page 243).

[o] A power calculation is an estimate of the size of the study population needed to detect an effect of a given size.

First and foremost, EPA provides no method for how it will determine the "adequacy" of the statistical power of a study on which to base its score, and provides no rationale for excluding studies with less than 80% statistical power. According to STROBE guideline developers, … "before a study is conducted power calculations are made with many assumptions that once a study is underway may be upended; further, power calculations are most often not reported" (32).

EPA's Metric 13 statistical power/sensitivity also appears to confuse bias with imprecision. Individual studies that are "underpowered" (for example, because in the real world the exposed population may not be large enough for statistical purposes even if they are health impacted) can still be potentially valuable to science-based decision-making. For example a small study may be imprecise but that should not be confused with whether it is biased (20); a small study can be imprecise but at the same time less biased than a larger study (17). Small "underpowered" studies can also be combined in a meta-analysis that increases the statistical power of the body of evidence to reflect the relationship between an exposure and a health impact. Additionally, "underpowered" studies that find a health effect to be present may be indicative of a larger effect size than anticipated. Thus, omitting such studies would severely bias the conclusions of the review.

Illustrative of how EPA's "analysis" metric could result in excluding high quality research that can inform science-based decision-making by EPA, in a 2017 systematic review by Lam et al. *"Developmental PBDE Exposure and IQ/ADHD in Childhood: A Systematic Review and Meta-analysis,"*(12) none of the 4 high-quality[p] studies included in the meta-analysis reported a power calculation, and yet together, these studies found "a 10-fold increase (in other words, times 10) in PBDE exposure associated with a decrement of 3.70 IQ points (95% confidence interval:0.83,6.56)." It is also notable that one of the studies in the meta-analysis, Herbstman et al. 2010, (38) was assessed by the review authors to be "probably high risk of bias" for "Incomplete Outcome Data."[q] As such, this otherwise high quality study, i.e., all of the other domains were "definitely" or "probably" low risk of bias, would meet EPA's criteria for "unacceptable for use" based on STROBE reporting guideline #15, "Report numbers of outcome events or summary measures over time".[r]

In short, the *Lam et al* systematic review, using the best available scientific methods, found that a ubiquitous environmental contaminant is impacting human intelligence, a finding that was subsequently reviewed and endorsed by the National Academy of Sciences (19). Yet EPA's SR review framework would exclude crucial pieces of this body of evidence based on the Agency's inaccurate, non-science-based criteria for deeming studies 'unacceptable.' This is contrary to TSCA's mandate to use the best available science. [s]

- **"Level of exposure"** is equated with a "serious flaw".

[p] "High quality" defined as "definitely" or "probably" low or very low risk of bias (Figure 2a in the *Lam et al* paper) based on specific and detailed definitions of risk of bias established before the review was conducted.

[q] The authors of the systematic review rated the Herbstman 2010 study "probably high risk of bias" for "incomplete outcome data" based on the following rationale: "Concerns regarding missing outcome data at each follow-up time on almost half the cohort of 210 with cord blood PBDE measurements; no argument is presented that would invalidate the possibility of a selection bias (i.e., likelihood that outcome data is missing is related both to outcome status and exposure)."

[r] See Table H-8 "Serious Flaws that Would Make Epidemiological Studies Unacceptable for Use in the Hazard Assessment" under the "outcome assessment domain" "Outcome measurement or characterization" metric (page 232) which specified STROBE guideline #15 to assess this metric.

[s] 15 USC §2625 (h)

EPA's "exposure characterization" domain for human studies includes the level of exposure as a fatal flaw, stating: "For all study types: The **levels** of exposure are not sufficient or adequate (as defined above)[t] to detect an effect of exposure (Cooper et al., 2016)." Unlike human experimental studies, which are largely precluded for ethical reasons, human observational studies can only be based on what exposures actually occur in the real world. EPA offers no explanation of how one could know whether the levels would be "sufficient or adequate" enough to detect an effect. Given the vagaries of this metric, it could be reasonably anticipated that it would permit EPA to arbitrarily exclude quality research from its decision-making.

We recommend: EPA should not use a quantitative scoring method to assess quality in individual studies; it should not conflate study reporting with study quality; and it should not exclude otherwise quality research based on a single reporting or methodological limitation. Rather EPA should employ a scientifically valid method to assess risk of bias of individual studies.

[t] EPA "as defined above" is unclear, presumably "as defined above" refers to the definition of the domain in Table H-2 page 223, "Evaluation of exposure assessment methodology that includes consideration of methodological quality, sensitivity, and validation of the methods used, degree of variation in participants, and an established time order between exposure and outcome."

3. EPA's TSCA systematic review framework does not consider financial conflicts of interest as a potential source of bias in research.

As observed by the Deputy Editor (West) of JAMA in 2010, "the biggest threat to [scientific] integrity [is] financial conflicts of interest" (39). Yet EPA's systematic review framework is silent on how it will take into account this empirically documented influence on the results of scientific research. Underscoring this EPA SR framework deficiency is the fact that recent studies empirically document that industry sponsorship produces research that is favorable to the sponsor (40, 41). The influence of financial ties on research can be traced to a variety of types of biases, and this conflict of interest needs to be distinguished from non-financial interests in the research, which can also affect research (42).

The fact that funding source needs to be accounted for in some manner is empirically supported and not a subject of scientific debate; what scientists differ on is *how* to best address funding as a potential source of bias (43, 44); for example, whether funding source is assessed as a specific risk of bias domain (43) or considered at multiple points in the evaluation (20, 44). For example, funding source is recommended as a factor to consider when evaluating risk of bias of individual studies for selective reporting, and then again for evaluating the body of evidence for publication bias, (45) and/or to be considered as a potential factor to explain apparent inconsistency within a body of evidence (14).

A 2017 Cochrane systematic review of industry sponsorship and research outcome concluded … "industry sponsorship should be treated as bias-inducing and industry bias should be treated as a separate domain" (40). The National Academy of Sciences in its review of the EPA IRIS program's SR method found that "Funding sources should be considered in the risk-of-bias assessment conducted for systematic reviews that are part of an IRIS assessment (17)(p 79).

Notably, EPA's exclusion of consideration of funding source and other potential conflicts of interests is also internally inconsistent with EPA's own improper reliance on STROBE guidelines as quality measures: STROBE guidelines item #22 specified that "the source of funding and the role of funders, could be addressed in an appendix or in the methods section of the article" (32).

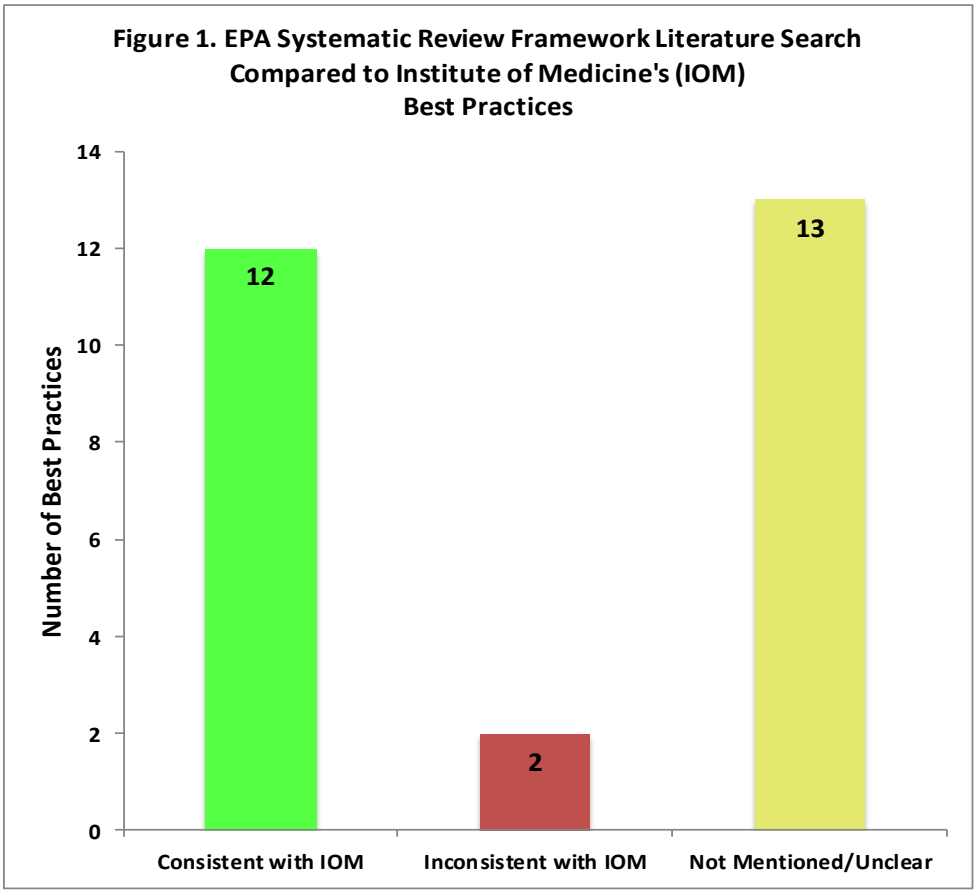
Importantly, including funding as a risk of bias as a domain does not mean excluding industry sponsored studies from EPA's hazard and risk assessment; it only means documenting funding as one of many domains of potential bias and evaluating its impact on the overall quality of the body of evidence.

We recommend: EPA should assess study and author funding source as a risk of bias domain for individual studies.

4. The literature review step of EPA's TSCA systematic review framework incorporates select best practices, but also falls short of, or is unclear about, many other best practices for conducting a systematic and transparent literature review.

Overall, we commend the EPA for its efforts to incorporate many best practices for a comprehensive literature search in its systematic review framework. We compared EPA's framework for systematic review to the Institute of Medicine's (IOM's) best practices for the literature review step of a systematic review (16)(See IOM 2011 Chapter 3. and TABLE E-1), which was applied by the National Academy of Sciences in its review of EPA's IRIS Program methods for systematic review (17)(See Table 4-1 Pp. 43-55).

We found EPA's framework to be consistent with 12 of IOM's 27 best practices for conducting a literature search (Figure 1 and Appendix 1). There are two key features of EPA's framework that are clearly inconsistent with IOM's best practices. EPA fails: (1) to include or exclude studies based on the protocol's pre-specified criteria, a practice that is critical to avoiding results-based decisions;[u] and (2) to use two or more members of the review team, working independently, to screen and select studies, which is an essential quality-assurance measure.[v]



[u] See our Comment #1 regarding the EPA framework's lack of a pre-defined protocol.

[v] EPA's framework, "Summary of the Title/Abstract Screening Conducted for the First Ten TSCA Risk Evaluations" (page 24) states that only one screener conducted the screening and categorization of titles and abstracts.

For the remaining 13 IOM best practices, EPA's framework is either unclearly stated (N=7) or the practice is not mentioned at all (N=6). However, based on the literature review methods presented in the First Ten TSCA Risk Evaluations, EPA's framework appears to have incorporated six additional best practices that are either unclear or not mentioned in EPA's SR framework: (1) work with a librarian or other information specialist trained in performing systematic reviews to plan the search strategy (IOM 3.1.1); (2) Design the search strategy to address each key research question (IOM 3.1.2); (3) Search regional bibliographic databases if other databases are unlikely to provide all relevant evidence (IOM 3.1.9); (4) Conduct a web search (IOM 3.2.5); and (5) Provide a line-by-line description of the search strategy, including the date of search for each database, web browser, etc. (IOM 3.4.1).

EPA should make its framework for conducting a literature review transparently congruent with all of IOM's best practices. This includes addressing two critical inconsistencies: (1) include or exclude studies based on the protocol's pre-specified criteria to prevent results-based decisions; and (2) Use two or more members of the review team, working independently, to screen and select studies, to ensure quality assurance. The transparency of the framework would be improved by specifying how EPA is addressing each best practice; at this juncture, how EPA intends to specifically handle many components of its literature searches could not readily be identified.

For example, the framework is unclear about whether EPA will include papers published in languages other than English. The exclusive reliance on English-language studies may lead to under-representation of the entire body of available evidence, and studies have also suggested that language bias might lead to erroneous conclusions (46). Furthermore, when considering the inclusion or update of an existing systematic review, studies have found that language-inclusive systematic reviews (including studies in languages other than English) were of the highest quality, compared with other types of reviews (47). Online translation tools are readily available to allow screeners to quickly evaluate study abstracts for relevance, and therefore we recommend EPA to incorporate non-English language studies in their screening and not simply exclude in advance these potentially relevant papers.

Additionally, EPA's framework should explicitly include rules for determining when the list of relevant studies will be considered final i.e., "stopping rules." Newer scientific studies will inevitably continue to appear in scientific journals and it will be impossible to continually attempt to include all these studies in a chemical assessment. To meet the deadlines as mandated by the Lautenberg Amendments, EPA should state clear stopping rules in the form of deadlines or criteria for when the body of included relevant studies will be finalized for the purposes of the chemicals assessment. We also strongly encourage EPA in its stated exploration of automation and machine learning tools,[w] which can help speed the production of EPA's systematic reviews.

We recommend: EPA should make its framework for conducting a literature review congruent with all of the Institute of Medicine's best practices, and explicitly include rules for when the list of relevant studies will be considered final.

[w] Footnote 9 page 23 states "In addition to using DistillerSR, EPA/OPPT is exploring automation and machine learning tools for data screening and prioritization activities (e.g., SWIFT-Review, SWIFT-Active Screener, Dragon, DocTER). SWIFT is an acronym for "Sciome Workbench for Interactive Computer-Facilitated Text-mining".

5. EPA's TSCA systematic review framework correctly recognizes that mechanistic data are not required for a hazard assessment, but EPA is not clear that these data, if available, can only be used to increase, and not to decrease, confidence in a body of evidence.

EPA's TSCA framework (page 172) states that EPA will use the evaluation strategies for animal and *in vitro* toxicity data to assess the quality of mechanistic and pharmacokinetic data supporting the model, and may tailor its criteria further to evaluate new approach methodologies (NAMs). We agree with EPA that mechanistic data need to be evaluated in a manner comparable to how other streams of evidence are evaluated. Data generated by alternative test methods (such as high-throughput screening methods) are not different than any other type of *in vitro* or cell-based assay data that would be considered in a systematic review. These kinds of assays provide mechanistic data. However, in this case, as described in comment # 2 above, EPA's use of its evaluation strategies for animal and *in vitro* toxicity data would entail using a quantitative scoring method that is incompatible with the best available science in fundamental ways. EPA should employ a scientifically valid method to assess risk of bias of individual studies in *all* streams of evidence, including mechanistic data.

EPA's framework (page 172) states, "the availability of a fully elucidated mode of action (MOA) or adverse outcome pathway (AOP) is not required to conduct the human health hazard assessment for a given chemical (emphasis added)." We strongly agree with EPA that mechanistic data are not needed for a hazard assessment. In addition, EPA's framework should be explicit that mechanistic data are only used to increase confidence in a hazard assessment, and never to decrease confidence.

The National Academy of Sciences explicitly considered how mechanistic data could be utilized in a systematic review for evidence integration (19). The committee came to two conclusions. First, the same protocol for evaluating relevance and study quality must be used with mechanistic data as for any other study. For example, in the report's case study on phthalates, the committee was not able to integrate results from high-throughput assays because the cell lines used were of unknown relevance to the *in vivo* mechanism of phthalate toxicity (19)(pg.78). Second, the foundation of the hazard classification in a systematic review is the animal and human data, with the mechanistic data playing a supporting role. If mechanistic data is relevant, it can be used to upgrade a hazard classification, or increase the confidence of a finding made based on evaluation of animal and human data. A hazard classification is never made based on high-throughput or other kinds of mechanistic data alone (19)(Pp. 158-9).

We recommend: EPA should be explicit that mechanistic data can only be used to upgrade a hazard classification, or increase the confidence of a finding made based on evaluation of animal and human data, and that these data will not be used to decrease confidence in a body of evidence.

6. EPA's TSCA systematic review framework is not independent of the regulatory end user of the review.

EPA's TSCA systematic review/risk assessment process is not independent of the TSCA risk management process, a conflict that is incompatible with best scientific methods. EPA's SR framework was developed and is being implemented by the Office of Chemical Safety and Pollution Prevention (OCSPP), which is also responsible for regulating the environmental exposures under TSCA review. In contrast, other EPA chemical assessment programs such as the IRIS program are intentionally placed in a non-regulatory research arm (the Office of Research and Development), to create separation from the Agency's program office responsible for regulatory decisions. This separation supports IRIS's ability to develop impartial chemical toxicity information independent of its ultimate use by EPA's program and regional office in risk assessment and risk management decisions. The National Academy of Sciences supported this in its 2018 report, stating that: "Current best practices [for systematic reviews in other medical disciplines] recommended by the Institute of Medicine (IOM 2011) suggest that the IRIS teams involved in the systematic-review process **should be independent of those involved in regulatory decision-making** who use the products of the systematic-review teams **(emphasis added)**" (15). This same principle should also be implemented across the Agency and specifically for TSCA assessments.

We recommend: EPA's systematic reviews should be produced independently of the regulatory end user of the review.

REFERENCES

1. Rennie D, Chalmers I. Assessing authority. JAMA. 2009;301(17):1819-21. Epub 2009/05/07. doi: 301/17/1819 [pii]10.1001/jama.2009.559. PubMed PMID: 19417202.
2. Fox DM. The Convergence of Science and Governance: Research, Health Policy, and American States. Berkeley, CA: University of California Press; 2010.
3. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA. 1992;268(2):240-8. Epub 1992/07/08. PubMed PMID: 1535110.
4. Woodruff TJ, Sutton P, The Navigation Guide Work Group. An Evidence-Based Medicine Methodology To Bridge The Gap Between Clinical And Environmental Health Sciences. Health Affairs. 2011;30(5):931-7. doi: 10.1377/hlthaff.2010.1219; PMCID: 21555477.
5. Woodruff TJ, Sutton P. The Navigation Guide sytematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. Environmenal Health Perspectives. 2014;122(10):A283.
6. Johnson PI, Sutton P, Atchley DS, Koustas E, Lam J, Sen S, Robinson KA, Axelrad DA, Woodruff TJ. The Navigation Guide - evidence-based medicine meets environmental health: systematic review of human evidence for PFOA effects on fetal growth. Environ Health Perspect. 2014;122(10):1028-39. Epub 2014/06/27. doi: 10.1289/ehp.1307893. PubMed PMID: 24968388; PMCID: 4181929.
7. Koustas E, Lam J, Sutton P, Johnson PI, Atchley DS, Sen S, Robinson KA, Axelrad DA, Woodruff TJ. The Navigation Guide - evidence-based medicine meets environmental health: systematic review of nonhuman evidence for PFOA effects on fetal growth. Environ Health Perspect. 2014;122(10):1015-27. Epub 2014/06/27. doi: 10.1289/ehp.1307177. PubMed PMID: 24968374; PMCID: 4181920.
8. Lam J, Koustas E, Sutton P, Johnson PI, Atchley DS, Sen S, Robinson KA, Axelrad DA, Woodruff TJ. The Navigation Guide - evidence-based medicine meets environmental health: integration of animal and human evidence for PFOA effects on fetal growth. Environ Health Perspect. 2014;122(10):1040-51. Epub 2014/06/27. doi: 10.1289/ehp.1307923. PubMed PMID: 24968389; PMCID: 4181930.
9. Vesterinen H, Johnson P, Atchley D, Sutton P, Lam J, Zlatnik M, Sen S, Woodruff T. The relationship between fetal growth and maternal glomerular filtration rate: a systematic review. J Maternal Fetal Neonatal Med. 2014:1-6. Epub Ahead of Print; PMCID: 25382561.
10. Johnson PI, Koustas E, Vesterinen HM, Sutton P, Atchley DS, Kim AN, Campbell M, Donald JM, Sen S, Bero L, Zeise L, Woodruff TJ. Application of the Navigation Guide systematic review methodology to the evidence for developmental and reproductive toxicity of triclosan. Environ Int. 2016;92-93:716-28. doi: 10.1016/j.envint.2016.03.009. PubMed PMID: 27156197.
11. Lam J, Sutton P, Halladay A, Davidson LI, Lawler C, Newschaffer CJ, Kalkbrenner A, Joseph J. Zilber School of Public Health, Windham GC, Daniels N, Sen S, Woodruff TJ. Applying the Navigation Guide Systematic Review Methodology Case Study #4: Association between Developmental Exposures to Ambient Air Pollution and Autism. PLoS One. 2016;21(11(9)). doi: 10.1371/journal.pone.0161851.

12. Lam J, Lanphear B, Bellinger D, Axelrad D, McPartland J, Sutton P, Davidson LI, Daniels N, Sen S, Woodruff TJ. Developmental PBDE exposure and IQ/ADHD in childhood: A systematic review and meta-analysis. Environmenal Health Perspectives. 2017;125(8). doi: doi: 10.1289/EHP1632.
13. Lam J, Koustas E, Sutton P, Cabana M., Whitaker E., Padula A, Vesterinen H, Daniels N, Woodruff TJ. Applying the Navigation Guide: Case Study #6. Association Between Formaldehyde Exposures and Asthma. In preparation. 2018.
14. National Toxicology Program. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. In: U.S. Department of Health and Human Services, editor.: Office of Health Assessment and Translation, Division of National Toxicology Program, National Institute of Environmental Health Sciences; 2015.
15. National Academies of Sciences, Engineering, and, Medicine. Progress Toward Transforming the Integrated Risk Information System (IRIS) Program: A 2018 Evaluation. Washington, D.C.: The National Academies Press; 2018.
16. Institute of Medicine. Finding What Works in Health Care. Standards for Systematic Review. Washington, D.C.: The National Academies Press.; 2011.
17. National Research Council. Review of EPA's Integrated Risk Information System (IRIS) Process. Washington, DC: National Academies Press; 2014.
18. National Research Council. Review of the Environmental Protection Agency's State-of-the-Science Evaluation of Nonmonotonic Dose–Response Relationships as They Apply to Endocrine Disruptors. Washington, DC: National Academies Press; 2014.
19. National Academies of Sciences, Engineering, and, Medicine. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, DC: 2017 2017. Report No.
20. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [Updated March 2011]: The Cochrane Collaboration. Available from http://www.cochrane-handbook.org.; 2011.
21. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology. 2011;64(4):383-94. doi: 10.1016/j.jclinepi.2010.04.026.
22. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW, Jr., Atkins D, Meerpohl J, Schünemann HJ. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). Journal of Clinical Epidemiology. 2011;64(4):407-15. doi: 10.1016/j.jclinepi.2010.07.017.
23. Vandenberg LN, Ågerstrand M, Beronius A, ., Beausoleild C, Bergman A, Bero LA, Bornehag C, Boyer CS, Cooper GS, Cotgreave I, Gee D, Grandjean P, Guyton KZ, Hass U, Heindel JJ, Jobling S, Kidd KA, Kortenkamp A, Macleod MR, Martin OV, Norinder U, Scheringer M, Thayer KA, Toppari J, Whaley P, Woodrufft TJ, Ruden C. A proposed framework for the systematic review and integrated assessment (SYRINA) of endocrine disrupting chemicals. Environment Health. 2016;In press.

24. Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC medical research methodology. 2014;14:43. Epub 2014/03/29. doi: 10.1186/1471-2288-14-43. PubMed PMID: 24667063.
25. Campbell Collaboration. Better evidence for a better world. 2018 [cited 2018 July 29]The Campbell Collaboration promotes positive social and economic change through the production and use of systematic reviews and other evidence synthesis for evidence-based policy and practice.]. Available from: https://campbellcollaboration.org/research-resources/writing-a-campbell-systematic-review.html
26. Herbison P, Hay-Smith J, Gillespie W. Adjustment of meta-analyses on the basis of quality scores should be abandoned. J Clin Epidemiol. 2006;59(12):1249-56. Epub 2006 Sep 11; PMCID: 17098567.
27. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA. 1999;282(11):1054-60. Epub 1999/09/24. doi: joc81641 [pii]. PubMed PMID: 10493204.
28. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol. 2010;25(9):603-5. Epub 2010 Jul 22. doi: 10.1007/s10654-010-9491-z.; PMCID: 20652370.
29. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. Biostatistics. 2001;2(4):463-71.
30. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, S. W. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. Control Clin Trials. 1995;16:62-73.
31. Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 510 [Updated March 2011]: The Cochrane Collaboration. Available from http://www.cochrane-handbook.org.; 2011.
32. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, Initiative. S. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Int J Surg. 2014;12(12):1500-24. doi: 10.1016/j.ijsu.2014.07.014.
33. Devereaux PJ, Choi PT, El-Dika S, Bhandari M, Montori VM, Schünemann HJ, Garg AX, Busse JW, Heels-Ansdell D, Ghali WA, Manns BJ, GH. G. An observational study found that authors of randomized controlled trials frequently use
concealment of randomization and blinding, despite the failure to report these methods. J Clin Epidemiol. 2004;57(12):1232-6; PMCID: 15617948.
34. Soares HP, Daniels S, Kumar A, Clarke M, Scott C, Swann S, B; D, Group. RTO. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. BMJ. 2004;328((7430)):22-4.; PMCID: PMC313900.
35. Lee W, Bindman J, Ford T, Glozier N, Moran P, Stewart R, M H. Bias in psychiatric case-control studies: literature survey. Br J Psychiatry. 2007;190:204-9.; PMCID: 17329739.
36. Kilkenny C, Browne W, Cuthill IC, Emerson M, Altman DG. Animal research: reporting in vivo experiments--the ARRIVE guidelines. J Cereb Blood Flow Metab. 2011;31(4):991-3. Epub 2011/01/06. doi: 10.1038/jcbfm.2010.220. PubMed PMID: 21206507; PMCID: 3070981.

37. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA, Group. P-P. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) :elaboration and explanation. BMJ. 2015;350:(g7647). doi: 10.1136/bmj.g7647.
38. Herbstman JB, Sjödin A, Kurzon M, Lederman SA, Jones RS, Rauh V, Needham LL, Tang D, Niedzwiecki M, Wang RY, Perera F. Prenatal exposure to PBDEs and neurodevelopment. Environ Health Perspect. 2010.
39. Rennie D. Integrity in scientific publishing. Health Serv Res. 2010;45(3):885-96. Epub 2010/03/27. doi: HESR1088 [pii]
10.1111/j.1475-6773.2010.01088.x. PubMed PMID: 20337732.
40. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev. 2017(2:MR000033.). doi: 10.1002/14651858.MR000033.pub3.; PMCID: 28207928.
41. White J, Bero LA. Corporate manipulation of research: Strategies are similar across five industries. . Stanford Law & Policy Review. 2010;21((1)):105-34. .
42. Bero L. Addressing Bias and Conflict of Interest Among Biomedical Researchers. JAMA. 2017;317(17):1723-4. doi: 10.1001/jama.2017.3854; PMCID: 28464166.
43. Bero L. Why the Cochrane risk of bias tool should include funding source as a standard item [editorial]. Cochrane Database Syst Rev. 2013;12:ED000075.
44. Sterne JA. Why the Cochrane risk of bias tool should not include funding source as a standard item. . Cochrane Database Syst Rev. 2013;(12)(:ED000076.). doi: 10.1002/14651858.ED000076.; PMCID: PMID: 24575440.
45. Viswanathan M, Ansari M, Berkman N, Chang S, Hartling L, McPheeters L, Santaguida P, Shamliyan T, Singh K, Tsertsvadze A, Treadwell J. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. 2012 AHRQ Publication No. 12-EHC047-EF.
46. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, Mierzwinski-Urban M, Clifford T, Hutton B, Rabb D. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. . International journal of technology assessment in health care. 2012;28((2)):138-44.
47. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. . Health Technol Assess. 2003;7((41)):1-90.